AFRL-RH-WP-TR-2009-0100100

Research Operations for Advanced Warfighter Interface Technologies

Timothy P. Ryan
General Dynamics AIS
Dayton OH 45431

June 2009

Final Report for January 2004 to June 2009

Approved for Public Release; distribution unlimited.

Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright-Patterson AFB OH 45433

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2009-0100100 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR

//signed//
Ronald E. Yates
Program Manager
Collaborative Interfaces Branch

//signed//
Daniel G. Goddard
Chief, Warfighter Interface Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
01-06-2009

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
January 2004 – June 2009

4. TITLE AND SUBTITLE
Research Operations for Advanced Warfighter Interface Technologies

5a. CONTRACT NUMBER
FA8650-04-C-6443

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Timothy P. Ryan

5d. PROJECT NUMBER
7184

5e. TASK NUMBER
HE

5f. WORK UNIT NUMBER
7184HE23

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
General Dynamics AIS
Dayton OH 45431

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright-Patterson AFB OH 45433-7022

10. SPONSOR/MONITOR'S ACRONYM(S)
711 HPW/RHCP

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RH-WP-TR-2009-0100

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
88 ABW/PA Cleared 13 Jul 2009; 88ABW-09-3056.

14. ABSTRACT
Execution of the Research Operations for Advance Warfighter Interface Technologies (ROADWIT) contract required a broad range of technical expertise supporting the mission of the Human Effectiveness Directorate (711 HPW/RHC), The mission of 711 HPW/RHC is to develop science and technology that improves the effectiveness of human-human and human-machine interaction within the Air Force. General Dynamics' mission, in support of 711 HPW/RHC research, is to ensure that the human operator is system enabling rather than system limiting thus resulting in the highest level of system effectiveness for the newest Warfighter technologies.

15. SUBJECT TERMS
Aerospace Vision, Bioacoustics, Command and Control Interfaces, Human Factors Engineering, Warfighter Interface

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UNCL	UNCL	UNCL	SAR	115	Ronald E. Yates
					19b. TELEPONE NUMBER (*Include area code*)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

This page intentionally left blank.

Table of Contents

List of Figures

List of Tables

Summary

Execution of the Research Operations for Advance Warfighter Interface Technologies (ROADWIT) contract required a broad range of technical expertise (Appendix A and B) supporting the mission of the Human Effectiveness Directorate (711 HPW/RHC), The mission of 711HPW/RHC is to develop science and technology that improves the effectiveness of human-human and human-machine interaction within the Air Force. General Dynamics' mission, in support of 711 HPW/RHC research, is to ensure that the human operator is system enabling rather than system limiting thus resulting in the highest level of system effectiveness for the newest Warfighter technologies. The result is Air Force systems that will continually surpass the capabilities of our adversaries thus discouraging them from directing hostile actions toward U.S. interests or reaping the consequences should they attempt to do so.

1. Introduction

This document provides a summary of work completed by General Dynamics under the work unit 71840871 Speech Interfaces for Multinational Collaboration for the period August 2004 to February 2009 under contract FA8650-04-C-6443. The next section describes how speech recognition systems were developed for 15 different languages, and presents three methods that were investigated for improving the performance of these systems. Section 3 describes how articulatory feature detectors were created for English and applied to speech recognition tasks in English, Russian, and Dari. Section 4 describes how speech synthesis systems were developed for 14 different languages, and provides a brief overview of two graphical user interfaces that were developed for creating new voices and synthesizing speech. Finally, section 5 summarizes the work completed and provides recommendations for future research.

2. Speech Recognition in 15 Languages

Speech recognition systems were developed for 15 different languages using the Hidden Markov Model (HMM) ToolKit (HTK). This section discusses these recognition systems and presents three methods that were investigated to improve the performance of these systems: Vocal Tract Length Normalization (VTLN), Speaker Adaptive Training (SAT), and the Recognizer Output Voting Error Reduction (ROVER) technique. Section 2.1 provides an overview of the baseline recognition systems developed for each language. Section 2.2 discusses VTLN and presents results obtained on English, Mandarin, and Russian. Section 2.3 provides an overview of SAT and presents results obtained on Russian and Dari. Lastly, Section 2.4 describes the ROVER technique.

2.1. Baseline Recognition Systems

This section discusses the baseline speech recognition systems that were developed for Arabic, Croatian, Dari, English, French, German, Japanese, Korean, Mandarin, Pashto, Russian, Spanish, Tagalog, Turkish, and Urdu. A total of seven different corpora were used to obtain coverage of all 15 languages, including the Topic Detection and Tracking (TDT4) Multilingual Broadcast News corpus [1], Phase II of the Wall Street Journal (WSJ1) corpus [2], CALLHOME Mandarin Chinese [3], HUB4 Mandarin Broadcast News Speech [4], GlobalPhone [5], the Language And Speech Exploitation Resources (LASER) Advanced Concept Technology Demonstration corpus, and the ARL Dari corpus. The TDT4, WSJ1, CALLHOME, and HUB4 corpora are available from the Linguistic Data Consortium, and the ARL Dari corpus was collected by Army Research Laboratory with support from AFRL. Table 1 lists the corpora used for each language, the speaking style of each corpus, the total amount of training data used to develop the recognizers, and the vocabulary size.

HMM-based recognition systems were trained for each language using HTK [6].[1] The feature set consisted of 12 Mel-Frequency Cepstral Coefficients (MFCCs), with cepstral mean subtraction, plus an energy feature. Delta, and acceleration coefficients were also included to form a 39 dimensional feature set. The acoustic models were state-clustered cross-word triphones. All HMMs included three states, with diagonal covariance matrices, and the state clustering was performed using a decision tree. An average of 16 mixture components were used for each HMM state.

1 Available at http://htk.eng.cam.ac.uk

Table 1: Overview of corpora

Language	Corpus	Speaking Style	Hours	Vocabulary Size
Arabic	TDT4	Broadcast News	37	47k
Croatian	GlobalPhone	Read	12	22k
Dari	ARL	Read	20	2k
English	WSJ1	Read	18	10k
French	GlobalPhone	Read	20	21k
German	GlobalPhone	Read	14	23k
Japanese	GlobalPhone	Read	26	18k
Korean	GlobalPhone	Read	16	50k
Mandarin	CALLHOME	Conversational	26	8k
Mandarin	HUB4	Broadcast News	30	18k
Pashto	LASER	Read	17	6k
Russian	GlobalPhone	Read	18	29k
Spanish	GlobalPhone	Read	17	19k
Tagalog	LASER	Read	9	5k
Turkish	GlobalPhone	Read	13	15k
Urdu	LASER	Read	45	8k

Trigram Language Models (LMs) were created for each language using the Carnegie Mellon University (CMU) Toolkit [7].[2] The LM probabilities were estimated using the train partition of each language, but the vocabulary was expanded to include all words in the corpus. Decoding was performed using both the HTK decoder HDecode and the Julius decoder [8].[3] The Word Error Rates (WERs) for each language are shown in Figure 1. HDecode yielded better performance than Julius in all languages.

2.2. Vocal Tract Length Normalization

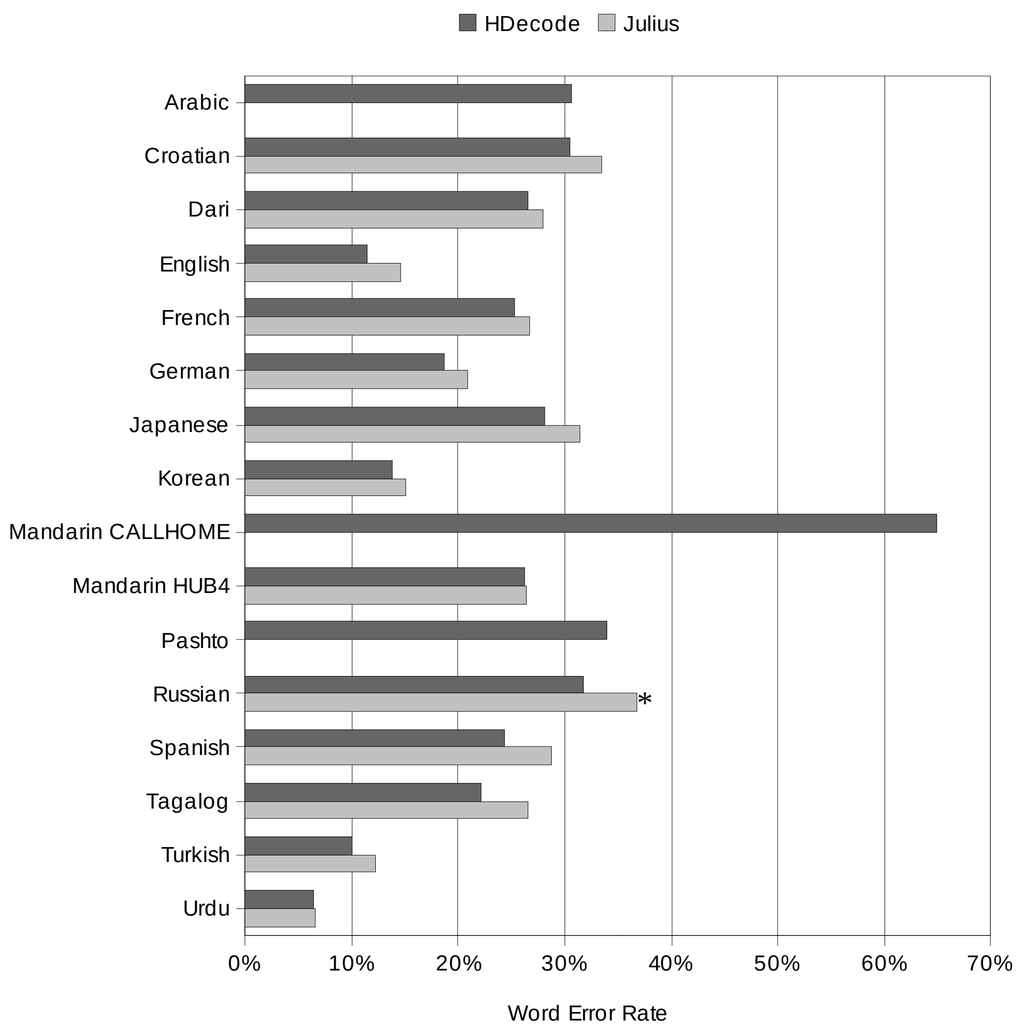
Vocal Tract Length Normalization (VTLN) attempts to compensate for different vocal tract lengths by linearly warping the frequency axis when performing filterbank analysis. Warping factors α for each speaker in the training set were selected using the following procedure [9]. First, single-mixture monophone HMMs with non-normalized MFCC

---

2 Available at http://www.speech.cs.cmu.edu

3 Available at http://julius.sourceforge.jp

features[4] were estimated from the complete training set of all speakers. Next, each utterance was phonemically aligned using the non-normalized HMMs and MFCC features computed using warping factors α=0.80,0.82,0.84,...,1.20. The value of α that gave the maximum score was selected for each speaker. Lastly, multiple-mixture triphone HMMs were estimated from the complete training set using the normalized MFCC features.



*

Figure 1: WER for each language (*Mandarin is expressed in character error rate).

The procedure used to select the warping factor α for each utterance in the test set can be summarized as follows. First, non-normalized multiple-mixture triphone HMMs with non-normalized MFCC features were used to hypothesize the word sequence for the utterance. Next, the utterance was phonemically aligned using the normalized single-mixture monophone HMMs and MFCC features computed using warping factors

4 Note that the term *normalization* is used to here to refer to MFCC features computed from a warped filterbank using α.

α=0.80,0.82,0.84,...,1.20. The value of α that gave the maximum score was selected for the utterance. Lastly, the normalized multiple-mixture triphone HMMs and normalized MFCC features were used to hypothesize the word sequence.

The VTLN procedure was evaluated on the WSJ1 English, CALLHOME Mandarin, and GlobalPhone Russian. The results for each language are shown in Table 2. Applying VTLN reduced the error rate by 1.0 percent on English, 1.7 percent on Mandarin, and 0.3 percent on Russian.

Table 2: WER for English and Russian, and character error rate for Mandarin.

Language	No VTLN	With VTLN
English	11.8%	10.8%
Mandarin	65.1%	63.4%
Russian	29.6%	29.3%

2.3. Speaker Adaptive Training

Speaker Adaptive Training (SAT) is a technique used to train Speaker Independent (SI) acoustic models that integrates speaker normalization as part of the model estimation procedure. The procedure used to implement SAT can be summarized as follows. First, multiple-mixture triphone HMMs were estimated from the complete training set of all speakers. Next, Constrained Maximum Likelihood Linear Regression (CMLLR)[5] was used to compute a set of linear transformations for each speaker. Lastly, the SI models were re-estimated using the speaker transforms to adapt the training features. This procedure was repeated three times to train the final model.

The decoding procedure can be summarized as follows. First, the original SI acoustic models were used to hypothesize the word sequence for each utterance. Next, each utterance was phonemically aligned using the SI acoustic models. These phoneme alignments were used to compute a single set of CMLLR transforms for each speaker using the SAT models. Lastly, the SAT models and CMLLR transforms were used to hypothesize the word sequence for each utterance. The SAT technique was evaluated on the GlobalPhone Russian and ARL Dari. The results are shown in Table 3. Applying SAT reduced the WER by 4.5 percent on Russian and 3.1 percent on Dari.

5 CMLLR is a feature adaptation technique that shifts the feature vectors such that each HMM state in the model is more likely to have generated the features.

Table 3: WER for Russian and Dari.

Language	No SAT	With SAT
Russian	29.6%	25.1%
Dari	26.6%	23.5%

2.4. ROVER

Recognizer Output Voting Error Reduction (ROVER) [10] is a technique for combining the hypothesized word sequences from multiple recognizers. The ROVER technique first aligns the word sequences output from the different recognizers and then selects the final word sequence according to the frequency of occurrence. This technique was evaluated on 12 different languages using the hypothesized word sequences from the HDecode, Julius, and SONIC [11] decoders. The SRover program from the University of Brno[6] was used to apply ROVER. Figure 2 shows the error rates obtained on each language. An improvement in system performance was obtained on all languages except English. Compared to the best individual system, the largest decrease in WER was 2.4 percent on French.

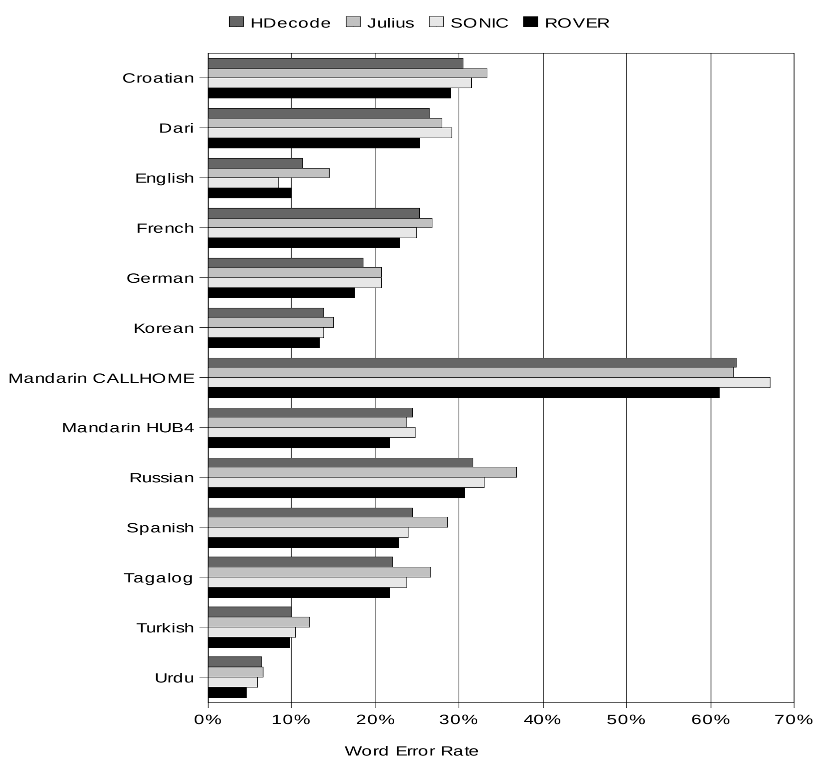


Figure 2: WER for each language using ROVER (*Mandarin is expressed in character error rate).

---

6 Available at http://speech.fit.vutbr.cz

3. Articulatory Feature Detection

Articulatory Features (AFs) describe the way in which speech sounds are produced. One of the most popular methods for classifying speech sounds using AFs is the International Phonetic Alphabet (IPA) [12]. Consonants are defined by AFs that describe the place of articulation, manner of articulation, and voicing status. Vowels are classified using AFs that describe both the tongue position and the shape of the lips. This chapter discusses two methods that were investigated for detecting English AFs. Section 3.1 describes how fusion-based AF detectors were created using Gaussian Mixture Models (GMMs) and two-class Multi-Layer Perceptrons (MLPs). Section 3.2 describes how multi-class MLPs were developed for English and incorporated into a Russian and Dari speech recognizer.

3.1. Fusion-based AF Detectors

This section discusses how fusion-based AF detectors were created for English and used in an HMM-based phoneme recognizer. Sections 3.1.1 and 3.1.2 describe how GMMs and MLPs were used to create AF detectors. Section 3.1.3 discusses two different procedures that were investigated for fusing the scores from the GMMs and MLPs, and presents results obtained on TIMIT. Lastly, Section 3.1.4 presents results obtained on the CSLU Multi-language Telephone corpus. Table 4 lists the AFs used to describe English speech sounds, with the exception of *silence (34)*, where the number in parenthesis indicates the feature number.

Table 4: AF for English consonants and vowels. Each AF is assigned a number.

CONSONANTS (0)	
Place	bilabial (1), labiodental (2), labialvelar (3), dental (4), alveolar (5), postalveolar (6), retroflex (7), palatal (8), velar (9), glottal (10)
Manner	plosive (11), nasal (12), tap or flap (13), fricative (14), approximant (15), lateral approximant (16), affricate (17)
Voicing	voiced (18), voiceless (19)

VOWELS (20)	
Tongue Height	close (21), near-close (22), mid (23), open-mid (24), near-open (25), open (26)
Tongue Fronting	front (27), near-front (28), central (29), near-back (30), back (31)
Lip Shape	rounded (32), unrounded (33)

3.1.1. GMM AF Detectors

GMM-based AF detectors were trained on the WSJ1 corpus using the GMM software package from MIT Lincoln Laboratory [13]. For each AF, a GMM was trained using frames where the feature was present, and a second GMM was trained using frames where the feature was absent. All models used 256 mixture components with diagonal covariance matrices. The feature set consisted of 12 MFCCs, with cepstral mean subtraction, plus an energy feature. Delta and acceleration coefficients were also included to form a 39 dimensional feature vector.

The scores for each AF were calculated as follows. Denote the presence of an AF as *f* and the absence of an AF as *g*. If we consider the speech feature vector *x*, then

$$\log \frac{p(f \mid x)}{p(g \mid x)} = \log p(x \mid f) - \log p(x \mid g) + \log p(f) - \log p(g)$$

The probabilities $p(x|f)$ and $p(x|g)$ were calculated from the feature-present and feature-absent GMMs, respectively. The probabilities $p(f)$ and $p(g)$ were estimated from the training data by counting the occurrences of each AF.

3.1.2. MLP AF Detectors

MLP-based AF detectors were trained on the WSJ1 corpus using the ICSI QuickNet software package.[7] A three-layered MLP (input: 39 units, hidden: 100 units, output: 2 units) was used to model each AF. The same MFCC feature set described in Section 3.1.1 was used as the input, and sigmoid activation functions were used on the hidden layer. The softmax function was used as the output activation function during training, however, it was removed when scoring the MLPs so that the outputs more closely approximated a Gaussian distribution. The final score for each AF was calculated by subtracting the output of the absent unit from the output of the present unit.

3.1.3. Score Fusion on TIMIT

This section describes two procedures that were investigated for fusing the scores from the GMM- and MLP-based AF detectors [14]. Both methods trained a fusion MLP for each AF to combine the scores. All fusion MLPs were trained on the TIMIT corpus [15]. *Fusion-1* combined the scores from the GMM- and MLP-based AF detectors for a given AF to form the final score for that AF. For example, the fusion MLP for the AF *plosive* used input features consisting of the output of the GMM-based plosive detector and the MLP-based plosive detector. *Fusion-2* combined the scores from all of the GMM- and

7 Available at http://www.icsi.berkeley.edu/Speech/qn.html

MLP-based AF detectors to form the final score for each AF; thus, the fusion MLP for each AF was provided information about all AFs from two different classifiers.

All fusion MLPs included 100 hidden units with sigmoid activation functions, and used the softmax output activation function for training. The fusion MLPs included a context window of nine; that is, the MLPs used the vectors at times $t\text{-}4, t\text{-}3, \cdots, t+3, t+4$ as input to classify the vector at time t. As in Section 3.1.2, the output activation function was removed prior to scoring and the score for the AF was calculated by subtracting the output of the absent unit from the output of the present unit.

Figure 3 shows the AF detection results obtained on the TIMIT test set. Each symbol represents the average Equal Error Rate (EER) of the individual detectors for the AF groups shown in Table 4. For the place and manner classifiers, the GMM-based detectors outperformed the MLP-based detectors; for all other groups the MLPs yield lower EERs. Fusion-1 yielded an average decrease in EER of 4.7 percent absolute compared to the best GMM- or MLP-based detector.[8] The best overall performance was obtained using the Fusion-2 procedure, which yielded an average decrease in EER of 8.2 percent absolute compared to the best GMM- or MLP-based detector.

The scores from the different AF detectors were used to form the feature set for an HMM-based phoneme recognizer. First, a vector was formed using the scores from the individual AF detectors.

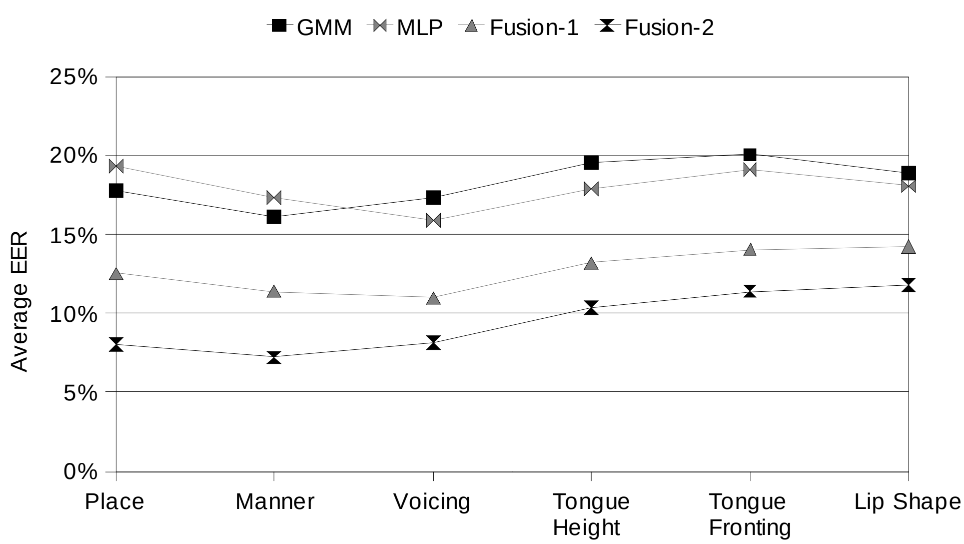


Figure 3: Average EER of the AF detectors on the TIMIT test set.

8 The term *best* is used here to refer to the detector with the minimum EER for each AF.

Next, these feature vectors were processed with a Karhunen-Loéve Transformation (KLT) that was estimated on the TIMIT train set. The KLT was included to decorrelate the individual AF scores so that diagonal covariance matrices could be used in the HMMs. Lastly, delta features were appended. Monophone and triphone HMMs were created for each feature set. All systems used three state HMMs with 16 mixtures per state and diagonal covariance matrices. Decoding was performed using a bigram phoneme LM that was estimated from the TIMIT train set using the CMU Toolkit. The MFCC feature set described in Section 3.1.1 was used for the baseline system.

Table 5 shows the Phoneme Error Rate (PER) obtained with each feature set on the TIMIT test set. The features created using the scores from the GMM-based detectors yielded the worst performance. An improvement in recognition performance was obtained using the scores from the MLP-based detectors, however, the PER was still higher than that of the baseline MFCC system. The Fusion-1 features outperformed both the GMM and MLP features sets, although an increase in performance over the baseline MFCC system was only obtained with monophone models. The best performance was obtained using the Fusion-2 features.

It is worth noting that the Fusion-2 monophone system yielded comparable performance to the MFCC triphone system. The option of using monophone instead of triphone models with the Fusion-2 features can be a significant advantage in terms of decoding time. Excluding the time required for feature extraction, decoding with each triphone system took approximately 750 minutes, whereas decoding with monophones was completed in about 20 minutes.

Table 5: PER obtained on the TIMIT test set.

	MFCC	GMM	MLP	Fusion-1	Fusion-2
Monophones	39.5%	42.1%	39.9%	38.8%	35.8%
Triphones	35.9%	40.8%	38.4%	38.4%	35.6%

3.1.4. Score Fusion on CSLU

This section discusses AF detection on the CSLU Multi-Language corpus [16]. Whereas TIMIT consists of lab-quality recordings of read speech with broad phonetic coverage, the CSLU corpus includes spontaneous telephone speech. Thus, these corpora differ in speaking style (read vs. spontaneous), channel type (close-talking microphone vs. telephone), balance of phonetic coverage, and sampling rate.

The WSJ1l and TIMIT corpora were first downsampled to 8 kHz and a second set of Fusion-2 AF detectors were retrained. Next, a set of Fusion-2 AF detectors were trained on the CSLU corpus. All AF detectors were created using the same procedure described in Sections 3.1.1-3.1.3. It should be emphasized that all fusion MLPs used scores from GMM- and MLP-based detectors trained on WSJ1 as input. Thus for the CSLU corpus, the base GMM- and MLP-based detectors were used for a different speaking style (read vs. spontaneous) and channel (close-talking microphone vs. telephone).

Figure 4 shows the EERs obtained with the Fusion-2 AF detectors. Each symbol type represents a different *train-test* combination. For example, TIMIT8-CSLU shows the detection performance obtained on the CSLU test set using Fusion-2 AF detectors trained on the TIMIT corpus downsampled to 8 kHz. The individual symbols represent the EER of each AF detector, where the feature numbers correspond to those given in Table 3.1. The best overall performance was obtained on the TIMIT8-TIMIT8 condition. The average EER across all AFs for this condition was 8.6 percent. When evaluated on the CSLU corpus, the fusion MLPs trained on TIMIT8 yielded an average EER of 14.1 percent, which is an increase of 5.5 percent compared to the results on TIMIT8. The average EER of the Fusion-2 AF detectors trained and evaluated on CSLU was 11.5 percent.

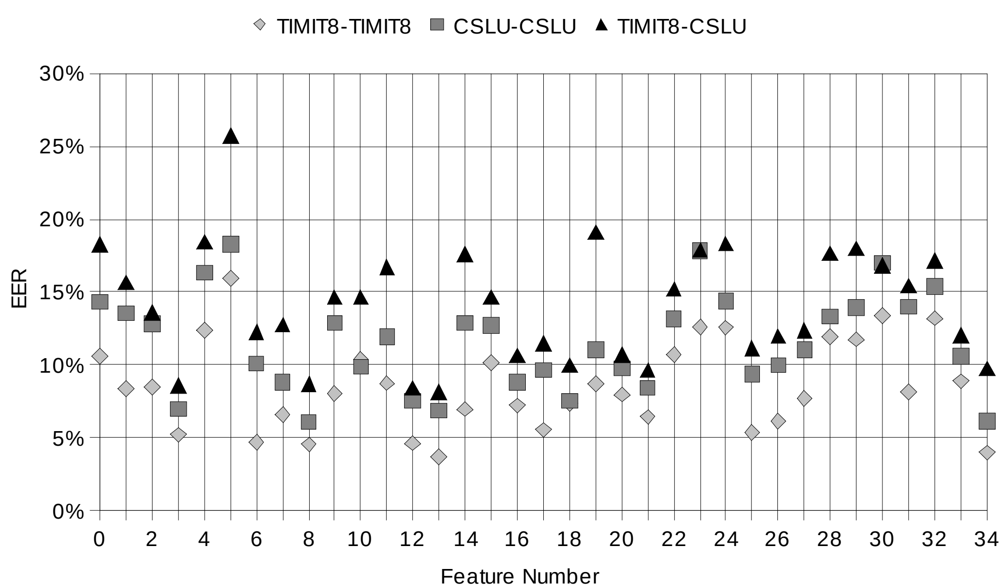


Figure 4: EER of the AF detectors on the CSLU test set.

From Figure 4 we can see that some of the AF detectors are more robust across both corpora than others. For example, the increase in EER on TIMIT8-CSLU compared to TIMIT8-TIMIT8 is less than 3.5 percent for the AFs *labialvelar (3), lateral approximant (16), voiced (18), vowel (20), close (21), near-back (30),* and *unrounded (33)*. The

increase in EER is greater than 8.0 percent for the AFs *alveolar (5), plosive (11), fricative (14)* and *voiceless (19)*. This suggests that certain AFs are less affected by speaking style and channel type than other AFs.

As in Section 3.1.3, the scores from the fusion MLPs were used to form the feature set for an HMM-based phoneme recognizer. Monophone and triphone HMMs were trained for each feature set on the CSLU corpus. The monophone models included 32 mixtures per state, and the triphone models included 12 mixtures per state. All systems used diagonal covariance matrices. Decoding was performed using a trigram phoneme LM that was estimated from the CSLU train partition using the CMU Toolkit. The MFCC feature set described in Section 3.1.1 was used for the baseline system.

Table 6 shows the PER obtained with each feature set on the CSLU test set. Both the TIMIT8 and CSLU Fusion-2 feature sets outperform the MFCC system. The best performance was obtained with the CSLU Fusion-2 features: compared to MFCCs, the PER was reduced by 2.0 percent absolute when decoding with either monophone or triphone models.

Table 6: PER obtained on the CSLU test set.

	MFCC	**TIMIT8 Fusion-2**	**CSLU Fusion-2**
Monophones	49.4%	48.6%	47.4%
Triphones	48.3%	47.4%	46.3%

3.2. AF Detection using Multi-Class MLPs

This section discusses how multi-class MLPs were used to create English AF detectors. Section 3.2.1 describes the procedure used to train the MLPs. Section 3.2.2 presents detection results obtained on SVitchboard and describes how the scores from the MLPs were used as the feature set for a speech recognizer. Lastly, Section 3.2.3 presents results obtained on Russian and Dari. Table 7 lists the features that were used to describe English speech sounds.

Table 7: Features used to describe English speech sounds [17].

Group	Feature Values
Place	alveolar, dental, labial, labiodental, lateral, none, postalveolar, rhotic, velar, silence
Degree	approximant, closure, flap, fricative, vowel, silence
Nasality	-, +, silence
Rounding	-, +, silence
Glottal State	aspirated, voiceless, voiced, silence
Vowel	aa, ae, ah, ao, aw1, aw2, ax, axr, ay1, ay2, eh, er, ey1, ey2, ih, iy, ix, ow1, ow2, oy1, oy2, uh, uw, none, silence
Height	high, low, mid, mid-high, mid-low, very-high, none, silence
Frontness	back, front, mid, mid-back, mid-front, none, silence

3.2.1. MLP AF Detectors

Two sets of MLPs were trained for each of the eight AF groups shown in Table 4. The first set used MFCCs as input, and the second set used Perceptual Linear Prediction (PLP) coefficients. The MFCC feature set was the same as described in section 3.1.1 except that both mean and variance normalization were applied on a per-conversation side basis. The PLP feature set included 12 PLP cepstral coefficients, plus energy, delta, and acceleration coefficients. As with the MFCCs, mean and variance normalization were also applied.

The MLPs were trained on the Fisher corpus [18,19] using the ICSI QuickNet software package. A context window of nine was used on the input layer, and the number of hidden units for each MLP was chosen using the same procedure as described in [17]. Sigmoid activation functions were used on the hidden layer. The number of output units for each MLP was set to the number of feature values for that AF group, and the softmax function was used as the output activation function.

3.2.2. AF Detection on SVitchboard

This section discusses AF detection on the SVitchboard corpus [20]. SVitchboard is a small vocabulary corpus that includes conversational telephone speech. A subset of 78 utterances include AF alignments that were manually produced. Figure 5 shows the frame level accuracy of the MLPs trained on Fisher using MFCC and PLP coefficients as input. For comparison purposes, the detectors from [17] were also evaluated on these

utterances. These detectors, referred to as *Frankel* in this document, use the same network typology and PLP feature set as the MLPs described in section 3.2.1. Overall, similar performance is obtained with each set of MLPs. The largest difference in accuracy is 2.0 percent (Frankel vs. PLP *degree*). The lowest accuracy was 75.8 percent (MFCC *place),* and the highest accuracy was 95.4 percent (Frankel *nasality).*

The scores from the MLPs were used to form the feature set for an HMM-based speech recognizer. First, a vector was formed using the scores from the individual AF detectors.

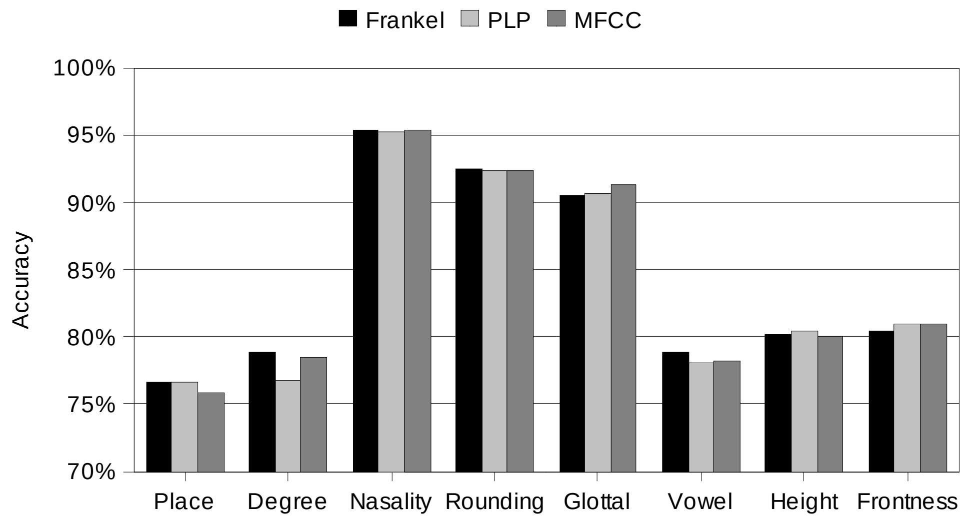


Figure 5: Frame level accuracy of the MLP-based AF detectors on the SVitchboard corpus.

When computing these scores, the output activation function was removed so that the scores more closely approximated a Gaussian. Next, these feature vectors were processed with a KLT that was estimated on the SVitchboard train set, and the top 26 dimensions were retained. This feature vector was appended to the PLP feature set described in Section 3.2.1 to form a 65 dimensional vector.

Within-word triphone HMMs were trained for each feature set. All systems used three state HMMs with 12 mixtures per state and diagonal covariance matrices. Decoding was performed using a bigram LM that was estimated from the SVitchboard train set using HTK. The PLP features formed the baseline system. Table 8 shows the WER obtained with each system. From Table 8 we can see that incorporating the scores from the MLPs yielded an improvement in system performance. The best WER was obtained with the PLP system that incorporated the Frankel MLPs: compared to the baseline PLP system, a reduction in WER of 6.0 percent was obtained. Note also that the MLP system with PLP input features yielded better performance than the MLP system with MFCC input features.

Table 8: WER on the SVitchboard 500 word vocabulary task.

Features	WER
PLP	50.6%
PLP + Frankel	44.6%
PLP + MLPs with PLP input	44.8%
PLP + MLPs with MFCC input	46.0%

3.2.3. Cross-Lingual AF Detection

The Frankel MLPs were also evaluated on the GlobalPhone Russian and ARL Dari. Whereas the Frankel MLPs were trained on English conversational telephone speech, the GlobalPhone Russian and ARL Dari corpora consist of read microphone speech. Thus, these corpora differ not only in language, but also in speaking style (conversational vs. read), channel type (telephone vs. microphone), and sampling rate.

The GlobalPhone Russian and ARL Dari corpora were first downsampled to 8 kHz and PLP features were extracted. These features were used as input to the Frankel MLPs, which were evaluated with the output activation functions removed. Next, a vector was formed using the scores from the individual AF detectors and processed with a KLT that was estimated on the train partition of each language. The top 26 dimensions were retained and appended to the MFCC feature set described in Section 2.1. This feature vector was used to train an HMM-based speech recognizer for each language. The HMM systems were trained using the same procedure described in Section 2.1 and decoding was performed using HDecode. The WER for each language is shown in Table 3.6. Incorporating the Frankel MLPs reduced the WER by 1.6 percent on Russian and 1.4 percent on Dari.

Table 9: WER on Russian and Dari.

Language	MFCC	MFCC + Frankel
Russian	29.6%	28.0%
Dari	26.4%	25.0%

4. Speech Synthesis in 14 Languages

Speech synthesis systems were developed for 14 different languages using the Hidden Markov Model (HMM) Speech Synthesis ToolKit (HTS). This section describes these systems and provides an overview of two different Graphical User Interfaces (GUIs) that were developed for creating new voices and synthesizing speech. Section 4.1 provides an overview of the baseline synthesis systems. Section 4.2 describes three English and two Urdu speech synthesis systems that were created using an expanded model set. Section 4.3 discusses the effect of modifying the Minimum Description Length (MDL) control factor. Section 4.4 discusses speaker clustering and adaptation for creating English and Mandarin voices. Lastly, Section 4.5 provides a brief overview of the GUIs that were developed.

4.1. Baseline Synthesis Systems

This section discusses the baseline synthesis systems that were developed for Arabic Iraqi, Croatian, Dari, English, French, German, Mandarin, Pashto, Russian, Spanish, Tagalog, Turkish, and Urdu. A total of six different corpora were used to obtain coverage of all languages, including the Spoken Language Communication and Translation System for Tactical Use (TRANSTAC) corpus, GlobalPhone, ARL, CMU Arctic [21], HUB4, and LASER. All of these corpora include speech data that were recorded with a 16 kHz sampling frequency. The CMU Arctic database was developed specifically for speech synthesis and includes automatically generated time-aligned transcriptions; all other corpora are only transcribed at the utterance level. Phoneme alignments for the TRANSTAC, GlobalPhone, ARL, HUB4, and LASER corpora were automatically generated using SONIC.

HMM-based speech synthesis systems were developed for each language using HTS-2.0 [22].[9] The feature set consisted of 25 Mel Cepstral Coefficients and the logarithm of the fundamental frequency (F0). Prior to computing the features, the DC mean was removed from each waveform file and amplitude normalization was applied to several of the corpora. The Mel Cepstral coefficients were calculated using the Speech Signal Processing ToolKit (SPTK),[10] and the F0 values were estimated using the ESPS method implemented in snack[11]. Delta and acceleration coefficients were also included to form a 78 dimensional feature vector.

9 Available at http://hts.sp.nitech.ac.jp

10 Available at http://sp-tk.sourceforge.net

11 Available at http://www.speech.kth.se/snack

Cross-word triphone Multi-Space probability Distribution (MSD)-HMMs [23] were trained for each language. All MSD-HMMs included five states with diagonal covariance matrices, and the state durations for each triphone were modeled by a Gaussian distribution. Decision tree based clustering was applied to the Mel Cepstrum, F0, and state duration distributions independently; thus, two decision trees were created for each MSD-HMM state, plus an additional decision tree for the state duration model. Table 10 lists the voices that were created for each language, the corpora used, the number of speakers used to train the voices, and the total amount of training data used to develop the synthesizers.

Table 10:List of Voices.

Language	Corpus	Voices	Speaker Count	Hours
Arabic Iraqi	TRANSTAC	Speaker1 Speaker2	370 30	10 3
Croatian	GlobalPhone	Male Female	32 48	5 7
Dari	ARL	Male1 Male2	15 15	2 2
English	CMU Arctic	Male Female SLT	4 2 1	3 2 1
French	GlobalPhone	Male Female	39 40	10 11
German	GlobalPhone	Male Female	60 5	13 1
Mandarin	HUB4	Male Wang Jianchuan Female Fang Jing	10 1 8 1	2 1 2 1
Mandarin	GlobalPhone	Male	15	4
Pashto	LASER	Random1 Random2	10 10	1 1
Russian	GlobalPhone	Male Female	49 44	9 9
Spanish	GlobalPhone	Male Female	38 46	8 10
Tagalog	LASER	Male Female	20 28	2 4
Turkish	GlobalPhone	Male Female	24 60	4 10
Urdu	LASER	Male Female	76 84	17 20

4.2. Full-Context Models

This section discusses the English and Urdu speech synthesizers that were created using an expanded model set. As mentioned in Section 4.1, the baseline synthesis systems for each language used cross-word triphone models. Although these models produce intelligible speech, there are numerous other contextual factors that can affect the overall prosody and naturalness of speech. In order to incorporate these contextual factors, the triphone labels for each speech database have to be expanded to include all features of interest. For example, the labels supplied with the HTS demos for the CMU Arctic database consist of 53 different contextual features, including syllable, accent, stress, part-of-speech, word, and phrase information. These labels are then used to define the acoustic models; thus, a separate MSD-HMM is trained for each phoneme that appears in a different context. Note that this can result in a very large model set prior to clustering. For example, the training data for the English SLT voice includes 38866 phoneme instances: using cross-word triphone labels requires 9480 unique MSD-HMMs, whereas using the expanded label set requires 38765 unique MSD-HMMs. An expanded set of labels were derived for Urdu that included syllable, word, and phrase information. These labels included a total of 31 different contextual features. Syllable information was explicitly marked in the pronunciation lexicon, and phrase information was derived by assigning a break wherever silence was labeled. Table 11 lists the expanded label set derived for Urdu.

Each of the three English voices and the two Urdu voices were retrained using the expanded labels. Overall, there was not a substantial improvement in voice quality. This may be due to the limited amount of speech data available to train different models for each phoneme in a particular context.

Table 11: Expanded label set for Urdu.

p1	the phoneme identity before the previous phoneme
p2	the previous phoneme identity
p3	the current phoneme identity
p4	the next phoneme identity
p5	the phoneme after the next phoneme identity
p6	position of the current phoneme in the current syllable (forward)
p7	position of the current phoneme in the current syllable (backward)
a1	the number of phonemes in the previous syllable
b1	the number of phonemes in the current syllable
b2	position of the current syllable in the current word (forward)
b3	position of the current syllable in the current word (backward)
b4	position of the current syllable in the current phrase (forward)
b5	position of the current syllable in the current phrase (backward)
b6	name of the vowel of the current syllable
c1	the number of phonemes in the next syllable
d1	the number of syllables in the previous word
e1	the number of syllables in the current word
e2	position of the current word in the current phrase (forward)
e3	position of the current word in the current phrase (backward)
f1	the number of syllables in the next word
g1	the number of syllables in the previous phrase
g2	the number of words in the previous phrase
h1	the number of syllables in the current phrase
h2	the number of words in the current phrase
h3	position of the current phrase in this utterance (forward)
h4	position of the current phrase in this utterance (backward)
i1	the number of syllables in the next phrase
i2	the number of words in the next phrase
j1	the number of syllables in this utterance
j2	the number of words in this utterance
j3	the number of phrases in this utterance

4.3. MDL Control Factor

Decision tree clustering in HTS is based on the MDL criterion [24]. The MDL criterion is used for selecting the questions when splitting nodes, and deciding when to stop growing the decision trees. A control factor λ is used to weight the penalty that the MDL criterion imposes for model complexity. As λ is increased, the penalty for a large model become larger and the stopping criterion is met sooner (thus producing a decision tree with fewer leaves). The English male and female voices described in Section 4.2 were retrained using λ = 1.0,0.7,0.4. The total number of leaves obtained for each λ are shown in Figure 6. As λ is increased, the total number of leaves for each of the decision trees decreases.

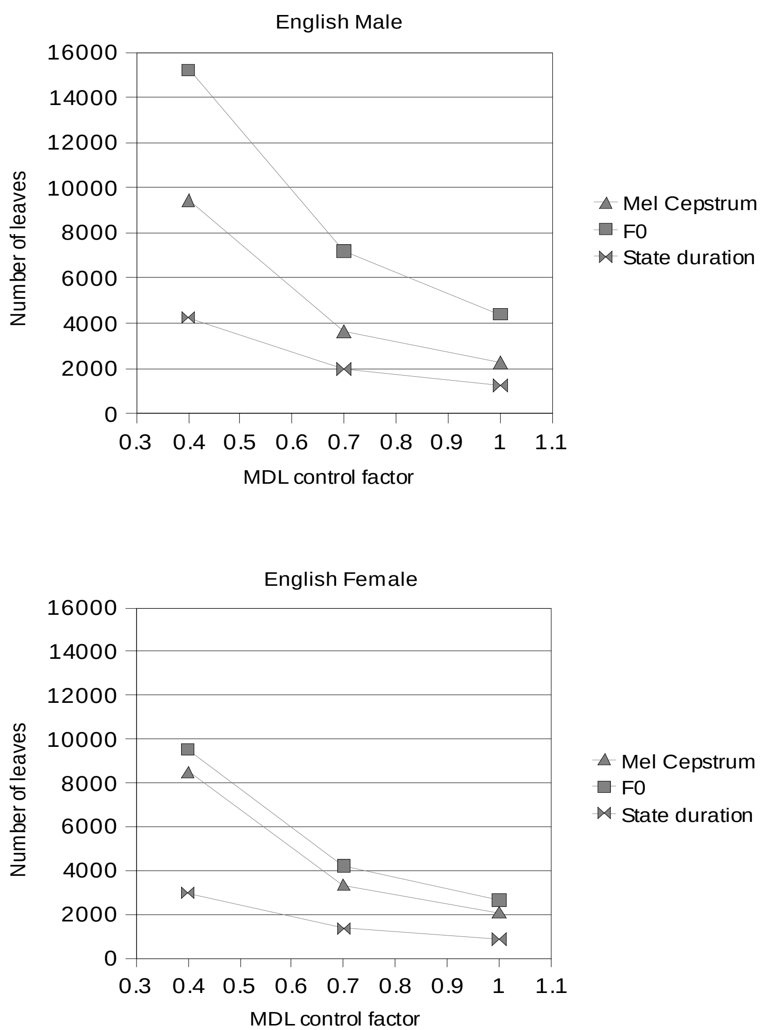


Figure 6: Total number of leaves generated for the English Male and Female voice when modifying the MDL control factor λ.

4.4. Speaker Clustering and Adaptation

This section discusses how speaker clustering and adaptation were used to create voices for Mandarin and English.[12] A total of 52 different Mandarin speech synthesis systems were trained on the GlobalPhone corpus using groups of three or more speakers. The speaker groups were defined based on the individual speakers F0 values and/or speaker recognition scores. Two additional voices were also created on the HUB4 Mandarin corpus by adapting the Male voice using speech from Wang Jianchuan, and adapting the Female voice using speech from Fang Jing. The adaptation transforms were estimated using Constrained Maximum Likelihood Linear Regression (CMLLR).

A total of 53 English speech synthesis systems were trained on Phase I of the Wall Street Journal (WSJ0) corpus [25] and WSJ1. These systems were developed using HTS-2.1.[13] Cross-word triphone MSD Hidden Semi-Markov Models (HSMMs) [26] were created for each voice using the same feature set as described in Section 4.1. As with the other corpora, the phoneme alignments were automatically generated using SONIC. The first 25 voices voices were created using groups of three or more speakers. The speaker groups were defined based on speaker recognition scores: 19 groups of speakers were derived from a speaker confusion matrix, and the remaining six groups were derived using a spectral clustering algorithm [27]. Next, one MSD-HSMM was trained using 3600 utterances from nine different speakers (~400 utterances from each speaker), and a second MSD-HSMM was trained using 3502 utterances from 20 different speakers (~200 utterances from each speaker). These models were adapted using speech from one of 22 different speakers to create the remaining 28 voices. Adaptation was performed using Constrained Structural Maximum-A-Posteriori Linear Regression (CSMAPLR), followed by MAP adaptation [28].

4.5. Synthesis GUIs

This section describes two GUIs that were developed for training and evaluating speech synthesizers. The first interface can be used to setup a speech synthesis experiment. This program allows the user to choose a set of speakers to train the voice and adjust system parameters related to speech analysis, model settings, and synthesis. Figure 7 shows two instances of the interface: the top one shows the speaker selection dialog, and the bottom one shows the spectrum analysis dialog. Once all configuration options have been specified, this program creates the makefiles for training and evaluating the system.

---

12 The speaker recognition experiments, F0 analysis, and speaker cluster definitions described in this section (except for those derived using the spectral clustering algorithm) were generated by Mr. Eric Hansen.

13 Available at http://hts.sp.nitech.ac.jp

The second interface can be used to synthesize speech, modify pronunciations, and create new voices by modifying the synthesis parameters. The text to synthesize can be entered using either the keyboard or read from a text file, and the pronunciations can be modified and saved on a per-speaker basis. The following synthesis parameters can be modified: all-pass constant, post-filtering coefficient, speech speed rate, multiplicative and additive constants for F0, voiced/unvoiced threshold, spectrum and F0 global variance weights, amplitude normalization constant, maximum state duration variance, and model interpolation coefficients. Figure 8 shows the main interface and pronunciation editor.

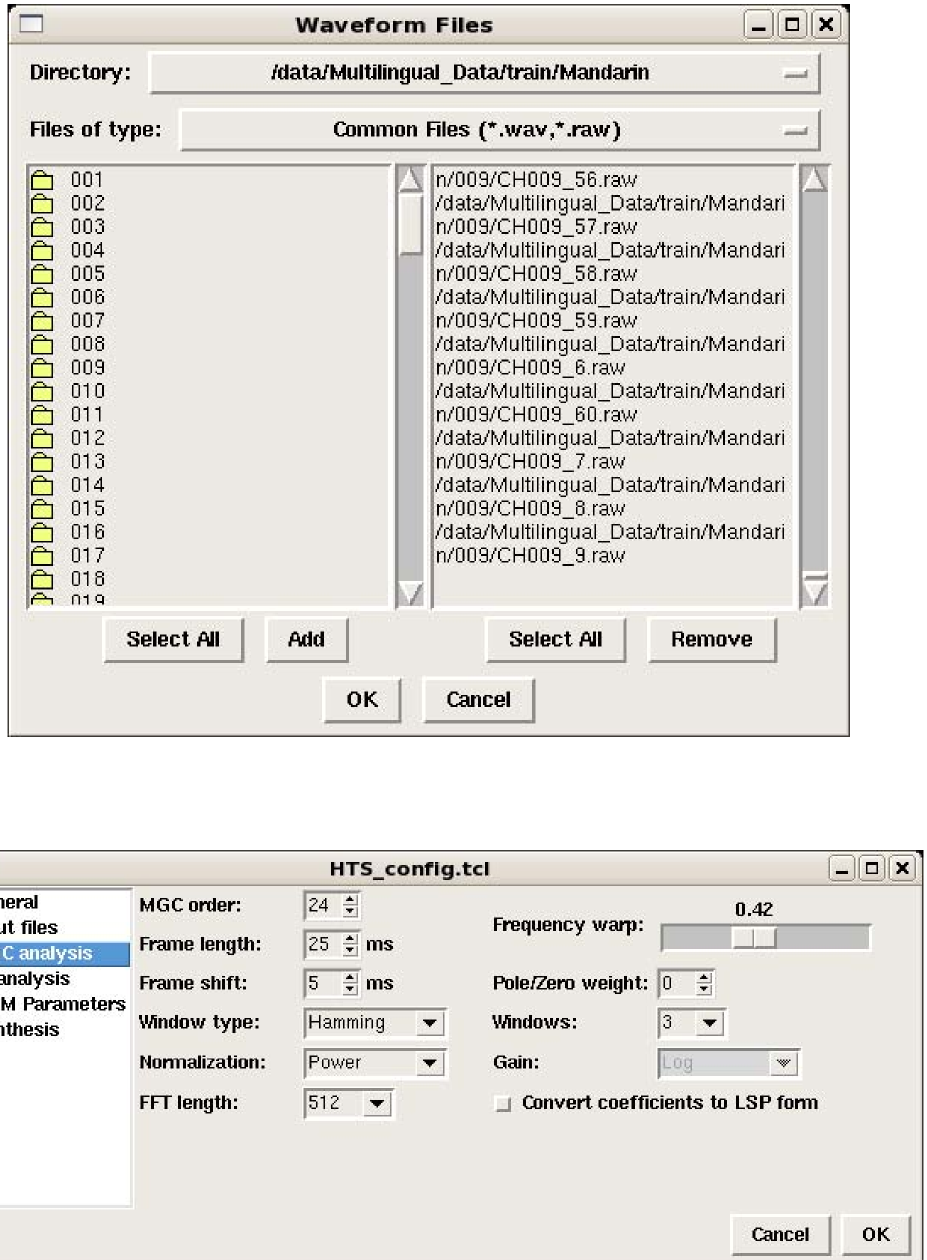


Figure 7: GUI for configuring a speech synthesis experiment. The speaker selection dialog is shown on top, and the spectrum analysis dialog is shown on the bottom.

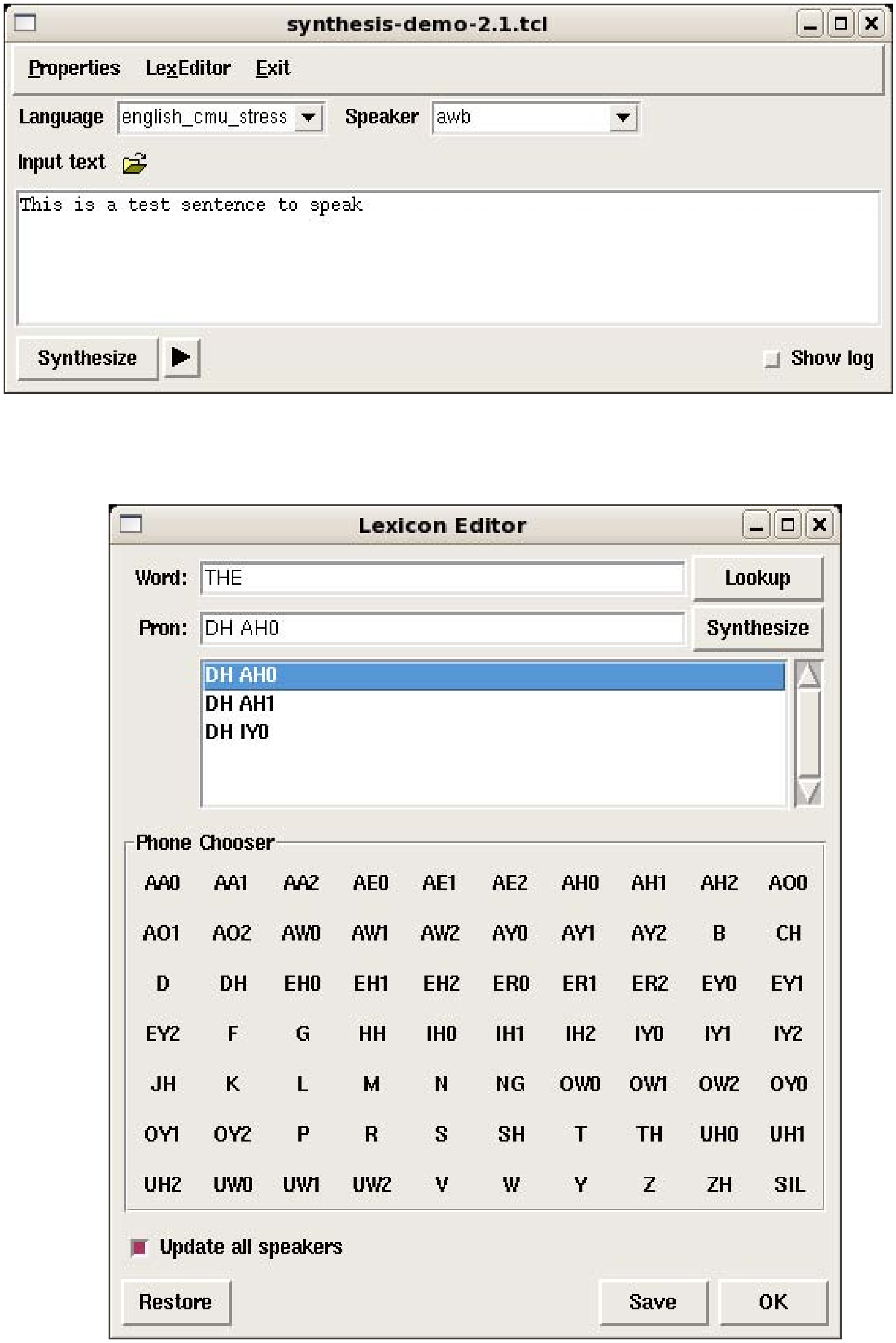


Figure 8: GUI for synthesizing speech. The main interface is shown on top, and the pronunciation editor is shown on bottom.

5. Summary and Recommendations

This document summarized work completed by General Dynamics during the period August 2004 to February 2009. Speech recognition systems were developed for 15 different languages using HTK. Three methods were investigated for improving the performance of these systems: VTLN, SAT, and the ROVER technique. Applying VTLN yielded improvements of 1.0 percent on English, 1.7 percent on Mandarin, and 0.3 percent on Russian. SAT reduced the WER by 4.5 percent on Russian and 3.1 percent on Dari. The ROVER technique yielded improvements in system performance of up to 2.4 percent. Given the substantial gains in system performance obtained with SAT, recommendations for future work include evaluating SAT across all languages, investigating how much speech data is needed from a single speaker to obtain an improvement in performance, and implementing an automatic method for detecting speaker changes and clustering speakers so that SAT can be applied to data where the speaker boundaries are unknown (*i.e.*, broadcast news).

AF detectors were developed for English using GMMs, two-class MLPs, fusion MLPs, and multi-class MLPs. The outputs of the detectors were used to form feature sets for HMM-based phoneme and word recognizers. On TIMIT, the Fusion-2 feature set yielded an improvement in PER of 3.7 percent compared to an MFCC system when decoding with monophones. On CSLU, the Fusion-2 features yielded improvements of 2.0 percent PER compared to MFCCs when decoding with either monophone or triphone models. On SVitchboard, appending the scores from the multi-class MLPs to PLP features yielded an improvement in WER of 6.0 percent. Finally, appending the scores from the English multi-class MLPs to MFCC features reduced the WER by 1.6 percent on Russian and 1.4 percent on Dari. Recommendations for future work include evaluating the English AF detectors across all languages, investigating methods for adapting the multi-class MLPs to different languages, and using alternative acoustic features for input to the MLPs.

Speech synthesis systems were developed for 14 different languages using HTS. Four methods were investigated for modifying these systems: expanding the model set to include additional contextual features, changing the MDL control factor, using speaker recognition scores and/or F0 values for grouping speakers to train voices, and applying speaker adaptation. Two GUIs were also developed for training and evaluating the speech synthesizers. Recommendations for future work include investigating how much speech data is needed to obtain an improvement when using an expanded model set, determining how much speech data is needed for speaker adaptation, and investigating the effects of using different speaker groupings to train the base model that is used for adaptation.

References

1. J. Kong and D. Graph, *TDT4 Multilingual Broadcast News Speech Corpus*, Linguistic Data Consortium, Philadelphia, 2005 (Available at http://www.ldc.upenn.edu).
2. *CSR-II (WSJ1) Complete*, Linguistic Data Consortium, Philadelphia, 1994 (Available at http://www.ldc.upenn.edu).
3. A. Canuan and G. Zipperlen, *CALLHOME Mandarin Chinese Speech*, Linguistic Data Consortium, Philadelphia, 1996 (Available at http://www.ldc.upenn.edu).
4. *1997 Mandarin Broadcast News Speech (HUB4-NE)*, Linguistic Data Consortium, Philadelphia, 1998 (Available at http://www.ldc.upenn.edu).
5. T. Schultz, "GlobalPhone: a Multilingual Speech and Text Database Developed at Karlsruhe University," in *Proc. of ICSLP*, Denver, Colorado, 2002.
6. Cambridge University Engineering Department, "The HTK Book," 2007 (Available at http://htk.eng.cam.ac.uk).
7. P. Clarkson and R. Rosenfeld, "Statistical Language Modeling using the CMU-Cambridge Toolkit," in *Proc. of ESCA Eurospeech*, Rhodes, Greece, September, 1997.
8. A. Lee, T. Kawahara, and K. Shikano, "Julius – an Open Source Real-Time Large Vocabulary Recognition Engine," in *Proc. of EUROSPEECH*, Aalborg, Denmark, September 2001.
9. L. Welling and H. Ney, "Speaker Adaptive Modeling by Vocal Tract Normalization," *IEEE Transactions on Speech and Audio Processing*, Vol. 10, No. 6, September 2002.
10. J. Fiscus, "A Post-Processing System to Yield Reduced Error Rates: Recognizer Output Voting Error Reduction (ROVER)," in *Proc. IEEE Workshop on Automatic Speech Recognition and Understanding*, Santa Barbara, California, December 1997.
11. B. Pellom, "SONIC: The University of Colorado Continuous Speech Recognizer," Technical Report TR-CSLR-2001-01, University of Colorado, March 2001.
12. International Phonetic Association, "Handbook of the International Phonetic Association," *Cambridge University Press*, 1999.
13. D. Reynolds, T. Quatieri, and R. Dunn, "Speaker Verification using Adapted Gaussian Mixture Model," *Digital Signal Processing*, Vol. 10, Nos. 1-3, January 2000.
14. B. Ore and R. Slyh, "Score Fusion for Articulatory Feature Detection," in *Proc. of Interspeech*, Antwerp, Belgium, August 2007.
15. J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, and N. Dahlgren, *TIMIT Acoustic Phonetic Continuous Speech Corpus*, Linguistic Data Consortium, Philadelphia, 1993 (Available at http://www.ldc.upenn.edu).

16. Y. Muthusamy, R. Cole, and B. Oshika, *CSLU: Multi-language Telephone Speech Version 1.2*, Linguistic Data Consortium, Philadelphia, 1993 (Available at http://www.ldc.upenn.edu).

17. J. Frankel, M. Magimai-Doss, S. King, K. Livescu, and Ö. Çetin, "Articulatory Feature Classifiers Trained on 2000 hours of Telephone Speech," in *Proc. of Interspeech*, Antwerp, Belgium, August 2007.

18. C. Cieri, D. Graff, O. Kimball, D. Miller, and K. Walker, *Fisher English Training Part I*, Linguistic Data Consortium, Philadelphia, 2004 (Available at http://www.ldc.upenn.edu).

19. C. Cieri, D. Graff, O. Kimball, D. Miller, and K. Walker, *Fisher English Training Part II*, Linguistic Data Consortium, Philadelphia, 2005 (Available at http://www.ldc.upenn.edu).

20. S. King, C. Bartels, and J. Bilmes, "SVitchboard 1: Small Vocabulary Tasks from SVitchboard 1", in *Proc. of Interspeech*, Libson, Portugal, 2005.

21. J. Kominek and A. Black, "CMU Arctic Databases for Speech Synthesis," Language Technologies Institute School of Computer Science, Carnegie Mellon University, CMU-LTI-03-177, 2003 (Available at http://festvox.org/cmu_arctic).

22. H. Zen, T. Nose, J. Yamagishi, S. Sako, T. Masuko, A. Black, and K. Tokuda, "The HMM-Based Speech Synthesis System Version 2.0," in *Proc. of ISCA SSW6*, Bonn, Germany, August 2007.

23. K. Tokuda, T. Mausko, and N. Miyazaki, "Multi-Space Probability Distribution HMM," *IEICE Transactions on Information and Systems*, Vol. E85-D, No. 3, March 2002.

24. Koichi Shinoda and Takao Watanabe, "Acoustic Modeling based on the MDL Principle for Speech Recognition," in *Proc. of Eurospeech*, Rhodes, Greece, September 1997.

25. J. Garofolo, D. Graff, D. Paul, and D. Pallett, *CSR-I (WSJ0) Complete*, Linguistic Data Consortium, Philadelphia, 2007 (Available at http://www.ldc.upenn.edu).

26. H. Zen, K. Tokuda, T. Masuko, T. Kobayashi, and T. Kitamura, "A Hidden Semi-Markov Model-Based Speech Synthesis System," I*EICE Transactions on Information and Systems*, Vol. E90-D, No. 5, May 2007.

27. U. Luxburg, "A Tutorial on Spectral Clustering," Technical Report TR-149, Max Planc Institute for Biological Cybernetics, August 2006.

28. J. Yamagishi, T. Kobayashi, Y. Nakano, K. Ogata, J. Isogai, "Analysis of Speaker Adaptation Algorithms for HMM-based Speech Synthesis and a Constrained SMAPLR Adaptation Algorithm," *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 17, No. 1, January 2009.

Appendix A

711 HPW/RHCP Support

Introduction

This report summarizes specific tasks completed by General Dynamics on 711 HPW/RHXS work unit 7184X07C, Crosslingual Audio Information Retrieval, for the period October 2005 to February 2009 under contract FA8650-04-C-6443.

The Air Force Research Laboratory's Speech and Communication Research, Engineering, Analysis, and Modeling (SCREAM) Laboratory has a commercially-available system to encode, index, archive, and search multimedia events such as news broadcasts. The system is from a company that was formerly called Virage, but which is now owned by a company called Autonomy. The Virage system contains a media encoder called a VideoLogger, and it has an audio indexing system from a company called BBN. The BBN audio indexing system gives the SCREAM Laboratory the capability to extract various metadata from audio and/or video content. The audio indexing system uses technologies such as automatic speech recognition (ASR), topic classification, speaker segmentation, speaker recognition, and named entity detection to extract information from audio. Specific information extracted includes spoken words, topic labels, identification of speakers, and entity tags such as person, location, organization, etc. The Virage system allows for the development of Media Analysis Plug-ins (MAPs), which can extend the media analysis capabilities of the VideoLogger.

This report discusses the development of a Virage MAP to allow for translating text generated by the ASR system as well as a plug-in that allows other ASR or audio processing systems to be integrated with the Virage system. Also discussed are the development of a search interface to allow for crosslingual audio information retrieval from foreign language media sources indexed by the Virage system as well as the collection of a corpus of foreign language materials to support the development of additional metadata detectors such as Interagency Language Roundtable (ILR) level, a United States Government-approved scale used to measure linguist proficiency level.[14]

An outline of this report is as follows. The next section describes the developed MAPs. Section 3 discusses the development of the search interface, while Section 4 describes the multilingual corpus collection. The final section summarizes the results and discusses future work.

---

[14] See `http://www.govtilr.org`

MAP Development

Overview

While the Virage and BBN products provide useful capabilities, researchers with the SCREAM Laboratory desire to extend and enhance these capabilities as well as to create similar solutions for other languages not currently supported. Two capabilities were developed to interface with the Virage VideoLogger. The SCREAM Virage Translator sends the words from an ASR system to an external system for language translation, and the SCREAM Virage Recognizer sends audio to an external system for processing by ASR, speaker recognition, and/or other signal processing.

SCREAM Virage Translator

The SCREAM Virage Translator uses a VideoLogger MAP, an utterance server, and an external machine translation (MT) engine to translate the VideoLogger "Words" text track. As words become available from the audio indexing system in near real-time, the plug-in sends the words, identified speakers, and timing information to the US. The utterance server groups words into sentence-like units, or utterances, based on the words, speakers, and timing information. Utterances are sent to an MT engine, and the translations are returned to the VideoLogger. Translation and utterance results are published to new VideoLogger text tracks called "Translation" and "Utterance." Figure A-1 shows the data flow for the SCREAM Virage Translator system.

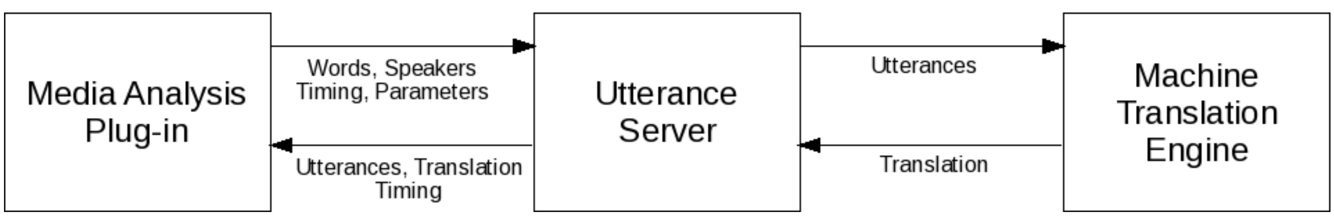


Figure A-1: SCREAM Virage Translator Data Flow

MAP Component

The MAP component of the SCREAM Virage Translator is a Microsoft Windows Dynamic-Link Library (DLL) developed using the Virage VideoLogger Software Development Kit (SDK) and Microsoft Visual Studio C++ 6. The plug-in monitors the "Words" and "Speakers" text track in the VideoLogger and sends the words, speakers, and associated timing information to the utterance server via a Transmission Control Protocol /Internet Protocol (TCP/IP) socket connection. The plug-in receives utterances, translations and associated timing information from the utterance server and publishes the data in the VideoLogger interface.

The plug-in has several configuration parameters:

29. Translation: This parameter identifies the translation to perform. Currently, this is limited to “Arabic to English” and “Chinese to English” based on the available BBN ASR systems integrated in the Virage system.
30. Server: This parameter identifies the hostname of the utterance server.
31. Port: This parameter identifies the TCP/IP port the utterance server listens on. The default value is 7890, but any valid TCP/IP port number is acceptable.
32. Intraword Delay (ms): This parameter is used by the utterance server to divide the running word sequence into utterances. To calculate utterances, words are accumulated until the speaker changes or the delay between any two subsequent words exceeds the Intraword Delay parameter. A typical value might be 750, which is also the default value.

The configuration parameters are set in the VideoLogger using the “Media Analysis” tab of the Preferences window as shown in Figure a-2. The Preferences window can be opened from the Options menu in the VideoLogger.

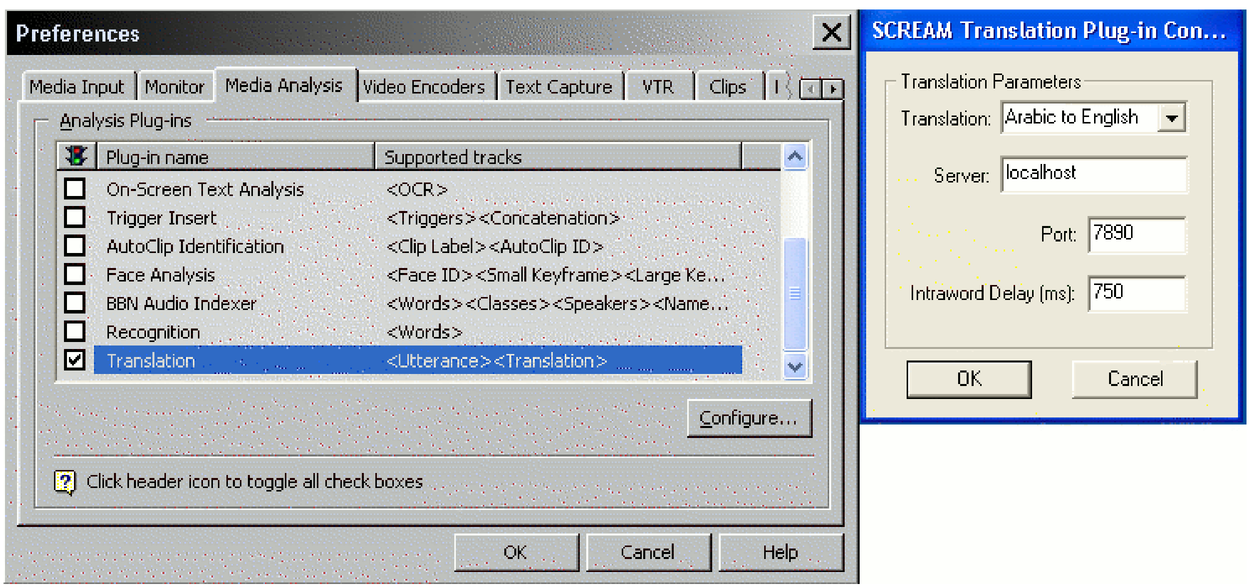


Figure A-2: MAP Configuration

If the Translation plug-in is enabled, the VideoLogger will send content from the “Words” and “Speaker” tracks to the utterance server identified by the configuration parameters. The Translation (e.g. “Arabic to English”) and the Intraword Delay parameters are sent as well. The VideoLogger will receive utterances and translations from the utterance server and display them in the VideoLogger as shown in Figure A-3.

Utterance Server Component

The utterance server is typically run on the same host as the VideoLogger; however, it is implemented in the Perl programming language and can run on any host that has Perl installed. The utterance server receives “Words” and “Speakers” and associated timing

from the VideoLogger. MT engines typically perform better when translating text with more context than they perform when just translating isolated words, so the utterance server is designed to collect words into sentence-like groups, or utterances. In order to determine the utterances, the server collects words that are from a particular speaker without long pauses. The length of a pause between words that will cause an utterance to end is the “Intraword Delay” as configured in the MAP component. Once a complete utterance is available, the utterance server connects to an MT engine to request a translation. The host providing the MT is configured near the top of the utterance server Perl script. The utterance server must connect to the MT engine on the appropriate port for the desired translation language pair. These ports are configured in a file called `ports`, which should be in the same directory as the utterance server.

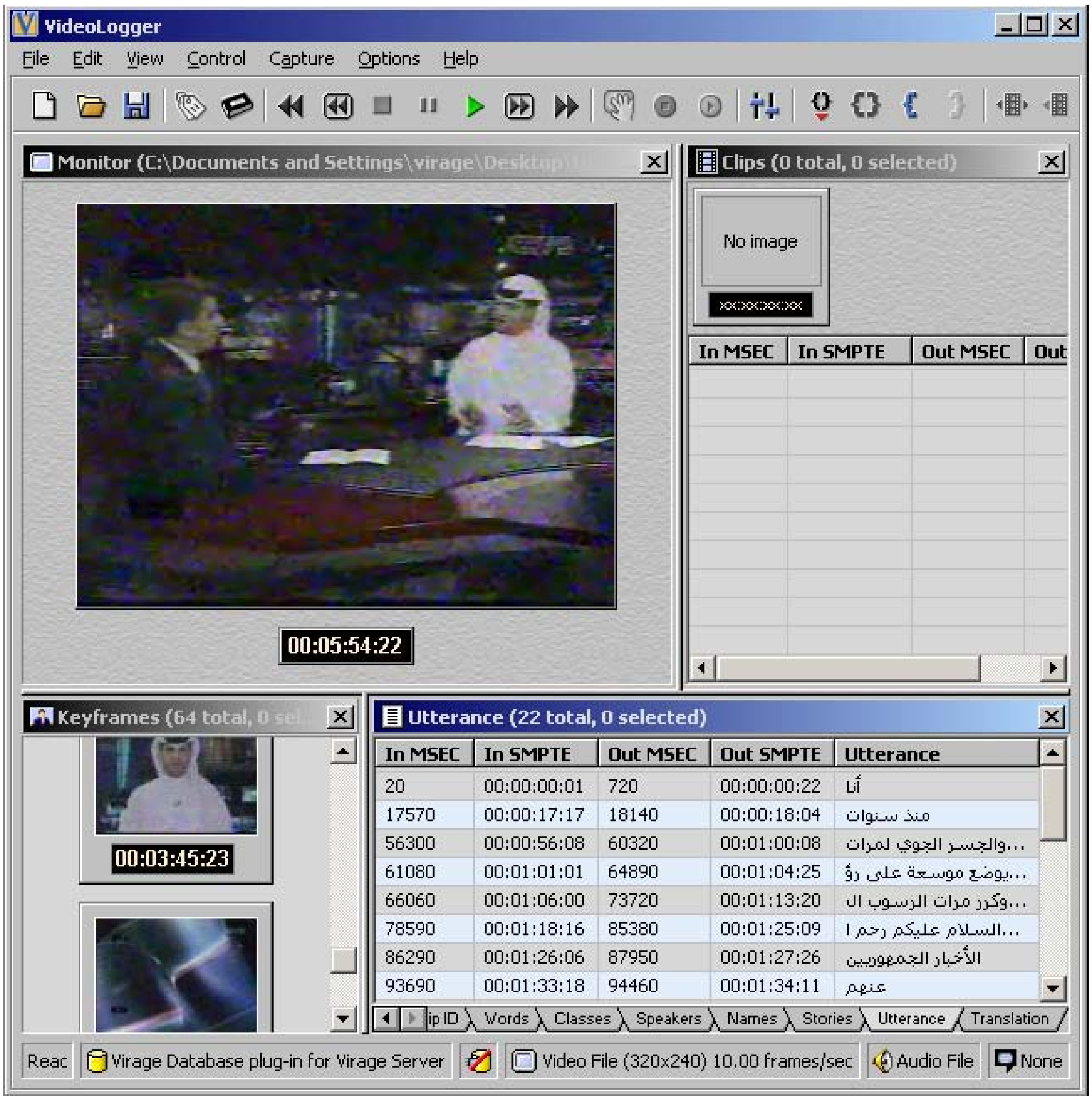


Figure A-3: VideoLogger with Utterance Track Displayed

The `ports` file is a list of language pairs, ports and descriptions as shown below:

```
ar_en 10036 Arabic to English
zh_en 20444 Chinese to English
```

The language pairs and port numbers must match the appropriate ports used by the MT engines. Currently, the utterance server only supports the "SYSTRAN simple text-based TCP/IP protocol," so the language pairs and port numbers shown in the example `ports` file are some of those supported by SYSTRAN.

The utterance server displays some log information as data is received and translated. Depending on the capabilities of the console or window, the foreign language characters may not display correctly. However, this does not affect the results displayed in the VideoLogger. Example log output from the utterance server is shown in Figure A-4.

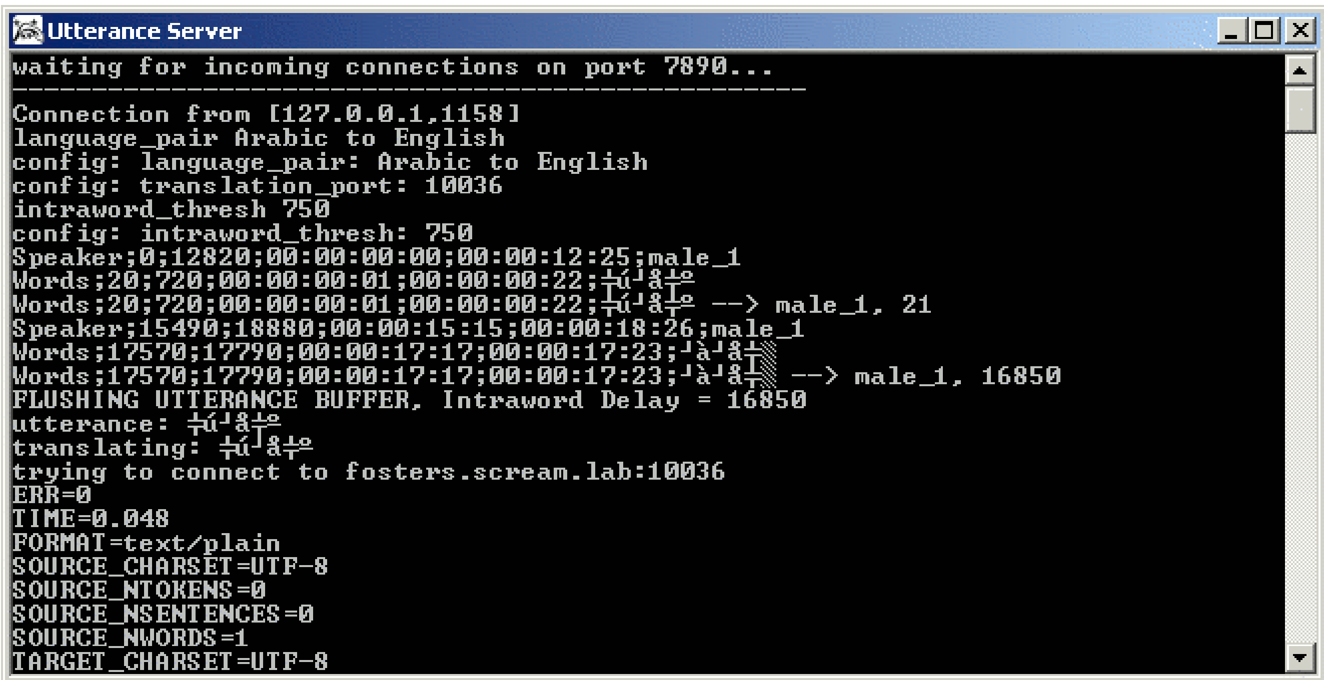


Figure A-4: Utterance Server Output

After each utterance is sent to the MT engine, the utterance server waits for the resulting translation. Once the translation is received from the MT engine, the utterance server sends the translation and the corresponding utterance back to the VideoLogger where they are displayed under the appropriate tabs.

Machine Translation (MT) Component

The SCREAM Virage Translator currently uses SYSTRAN MT engines—specifically, the SYSTRAN Version USG 4.2 engines hosted on Solaris 8. If another MT engine were used, the utterance server would require modifications to interface with the desired engine.

SCREAM Virage Recognizer

The SCREAM Virage Recognizer uses a VideoLogger MAP to interface with an audio processing component, such as an ASR system like SONIC, a large vocabulary continuous speech recognition system developed at the University of Colorado at Boulder [1, 2]. As the VideoLogger receives audio data from a multimedia event, the MAP sends a stream of audio data to an ASR server. After receiving enough data on which to perform ASR, the ASR server sends the recognized words back to the MAP. Figure A-5 shows the data flow for the SCREAM Virage Recognizer system.

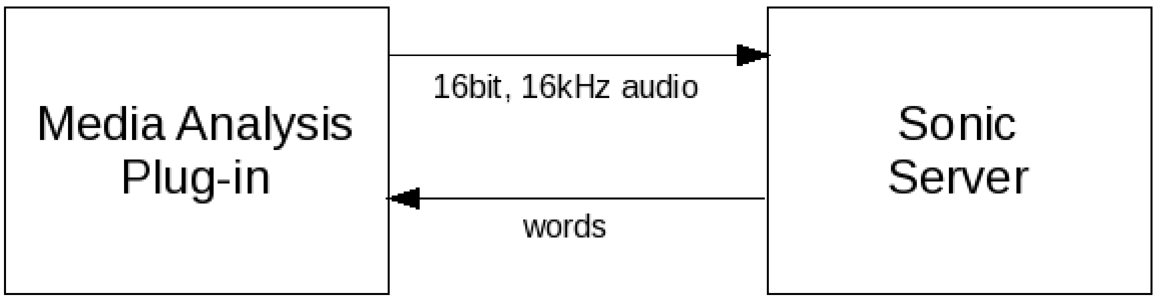


Figure A-5: SCREAM Virage Recognizer Data Flow

MAP Component

The MAP component of the SCREAM Virage Recognizer is a Microsoft Windows Dynamic-Link Library (DLL) developed using the Virage VideoLogger SDK and Microsoft Visual Studio C++ 6. The plug-in requests the raw audio signal from the VideoLogger. Because the SONIC Server requires audio sampled at 16 kHz, the plug-in resamples the audio from the native VideoLogger rate, 22 kHz, to 16 kHz. The resampled audio is sent to the SONIC Server over a TCP/IP socket connection. Words recognized by the SONIC Server are received by the plug-in and written to the VideoLogger's media-analysis log file.

Audio Resampling

The MAP uses libresample, a real-time library for sampling rate conversion by Dominic Mazzoni.[15] Libresample is free software released under the Lesser General Public License (LGPL) from the Free Software Foundation. When raw audio becomes available to the VideoLogger, the MAP uses libresample to change the audio sampling rate as necessary to interface with the SONIC Server.

[15] See http://ccrma.stanford.edu/~jos/resample/Free_Resampling_Software.html

Timing Problems

While developing the SCREAM Virage Recognizer plug-in, we encountered errors when interfacing with the SONIC Server whereby the timing information for the individual recognized words was not correct. As a result, the current version of the Recognizer plug-in writes the recognized words to the VideoLogger media-analysis log file rather than to a track. If this timing issue is resolved in the future, then SONIC could be fully integrated. Other ASR servers can be fully integrated as long as they return the correct starting and ending times for each word.

SONIC Server Component

The SONIC Server component is the SONIC recognizer running in `live_mode` using the following configuration (stored in the `sonic.cfg` file):

```
-langmod_file            kb/wsj-5k-cnp.bin
-dictionary              kb/wsj-5k.lex
-phone_config            kb/phoneset.cfg
-acoustic_mod            kb/wsj-i.mod
-filler_file             kb/wsj.filler
-filler_penalty          0.0
-word_entry_beam         80.0
-state_beam              160.0
-word_end_beam           80.0
-lm_scale                25.0
-rescore_lm_scale        25.0
-word_trans_penalty      -12.5
-state_dur_scale         2.5
-short_word_penalty      0.0
-sample_rate             16000.0
-max_active_states       40000
-auto_end_point          1
-end_point_padding       125
-max_word_ends           400
-confidence              1
-confidence_am_scale     25.0
-live_mode               1
-push_to_talk            0
```

The SONIC Server is started using the following command:

```
SONIC/2.0-beta5/bin/i686-Linux/SONIC_server -g -port 5555 -c SONIC.cfg
```

The server can be tested by sending it a test audio file with the following command:

```
SONIC/2.0-beta5/bin/i686-Linux/SONIC_client -h localhost -p 5555
test.raw
```

Media Search Interface

The rich metadata that results from the audio indexing performed by the Virage VideoLogger or similar systems can be useful for numerous applications. Crosslingual audio information retrieval to support language learning is one such application of interest to researchers in the SCREAM Laboratory. The media search interface application and multilingual corpus collection (discussed in the next section) were projects conducted to support the language learning application.

The SCREAM Media Search, Figure A-6, is a web based application to search the foreign language multimedia data collected, encoded, and indexed with the Virage VideoLogger or similar system. The media search application demonstrated a method of searching the metadata for specific keywords in English or the foreign languages supported by ASR systems and displaying the results with additional analysis data such as vocabulary coverage ranking.

Figure A-6: SCREAM Media Search

The media search engine was developed using the full-text search capabilities of mySQL[16]—namely, Boolean search, natural language search, and query expansion search. Full-text searching is performed using "MATCH() ... AGAINST" syntax. "MATCH()" takes a comma-separated list that names the columns to be searched. "AGAINST" takes a string to search for and an optional modifier that indicates what type of search to perform. The search string must be a literal string, not a variable or a column name.

[16] See `http://www.mysql.com`

A Boolean search interprets the search string using the rules of a special query language. The string contains the words to search for. It can also contain operators that specify requirements such that a word must be present or absent in matching rows, or that it should be weighted higher or lower than usual.

A natural language search interprets the search string as a phrase in natural human language (i.e., a phrase that could occur in free text); there are no special operators. However, a stopword list (i.e., a list of common words such as "the," "and," "a," and "an" that do not carry much information content for retrieval purposes) is applied, so that the presence or absence of the stopwords in the query or the database does not affect the search results. In addition, words that are present in 50 percent or more of the rows are considered common and do not match.

A query expansion search is a modification of a natural language search. The search string is used to perform a natural language search. Then, words from the most relevant rows returned by the search are added to the search string, and the search is performed again. The query returns the rows from the second search. A query expansion search can boost recall (i.e., the percentage of relevant documents that are returned) at a cost of lowering precision (i.e., the percentage of returned documents that are relevant).

The information in the database is searchable by keywords, but the search can be narrowed to search only particular tracks and languages via Track and Language parameters. Selectable tracks for searching include: Closed Caption, Names, Speakers, Speaker ID, Speech, Stories, Translation, Utterance, Words, and All Tracks.

The search return links to the original video streams according to the time-code values stored within the database. The videos are available to play as a full clip of the event or as a user-defined portion of the clip according to the search result values.

Multilingual Corpus Collection

The collection of a multilingual corpus was initiated for use in developing detectors for Interagency Language Roundtable (ILR) level as well as other metadata. The corpus was created by retrieving lessons from the Global Language Online Support System (GLOSS), [17] a web site provided by the Curriculum Development Division of the Defense Language Institute Foreign Language Center (DLIFLC). GLOSS language lessons are developed for students and Department of Defense linguists to support language learning and sustainment in reading and listening using authentic materials such as magazine articles, TV and radio broadcasts, and interviews.

[17] See `http://gloss.lingnet.org`

At the time of the corpus collection, the GLOSS site provided materials in 27 languages, grouped by ILR proficiency level, skill modality, competence, and topic. The ILR proficiency level for each lesson was labeled by trained raters according to ILR standards. The ILR scale consists of six “base levels” ranging from 0, No Proficiency, to 5, Functionally Native Proficiency, with intervening “plus levels” that indicate when the required proficiency level substantially exceeds one base skill level, but does not fully require the criteria for the next “base level.” The lessons retrieved from the GLOSS site consisted almost entirely of lessons rated in the 2, 2+, and 3 levels. The skill modality refers to whether the lesson is based on listening or reading. The competence refers to whether the lesson primarily focuses on lexical, discourse, structural, or socio-cultural content of the material. The topics covered in the lessons are: Culture, Economy, Environment, Geography, Military, Politics, Science, Security, Society, and Technology.

A list of available lessons was created for each language using the GLOSS naming schema, and the lists were used in a web scraping tool to collect the relevant files. Each lesson consisted of multiple HTML, image, and multimedia files. The collected HTML files were edited to pull out the source text and the English translations for further use. In total, the amount of captured information measured over two gigabytes with nearly 21,000 files.

Results and Future Work

Results

The SCREAM Virage Translator successfully integrates the Virage VideoLogger, BBN audio indexing system, and SYSTRAN MT engines to provide language translations of live or recorded multimedia events. Although the system currently only handles Arabic to English and Chinese to English translations, it could be easily extended to additional languages if the necessary ASR and MT capabilities were available.

The combination of the MAP and the utterance server was a design to keep the MAP minimalistic and increase flexibility. This flexibility could be enhanced by a more general version of the MAP that would allow any Virage VideoLogger text track to be retrieved instead of just the “Words” track.

The SCREAM Virage Recognizer successfully demonstrates the use of the SONIC recognizer to perform ASR on Virage VideoLogger media events. Development was not 100 percent completed after the timing problems with the interface to the SONIC Server were discovered. The ASR results from the SONIC Server were written to the VideoLogger media analysis log file instead of being published as a text track in the VideoLogger interface as publishing a text track in the VideoLogger requires a correct starting and ending time for each element.

A search interface was developed that allowed for crosslingual audio information retrieval based on the metadata in the Virage database. The search can be narrowed to search only particular tracks and languages via Track and Language parameters.

A multilingual corpus was collected to facilitate the development of detectors for ILR level and other metadata. If these detectors are integrated into the Virage system to provide additional metadata tracks, then these tracks can be made available to the search interface.

Future Work

One potential method of solving the SCREAM Virage Recognizer timing problem while still using the SONIC Server would involve sending segments of audio to the SONIC server via individual TCP/IP socket connections. The length of each audio segment could be used to calculate the starting and ending time for the group of words recognized for each segment. This method would also require the MAP component to implement a robust speech/silence detector to avoid segmenting the audio during active speech. Using

individual TCP/IP socket connections for each audio segment would also incur additional network overhead as a network socket would be opened and closed for each speech segment. While the SCREAM Virage Recognizer currently only communicates with the SONIC Server for ASR, it could be extended easily to support other ASR systems.

Future work by SCREAM Lab researchers will focus on developing various metadata detectors, such as detectors for ILR level. When these detectors are complete, they can be integrated into the Virage system with plug-ins and their associated metadata tracks can be provided to the search interface.

References

[1] B. Pellom, *SONIC: The University of Colorado Continuous Speech Recognizer*, University of Colorado, Technical Report TR-CSLR-2001-01, (Boulder, Colorado), March 2001.

[2] B. Pellom and K. Hacioglu, "Recent improvements in the CU SONIC ASR system for noisy speech: The SPINE task," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Hong Kong), April 2003.

Appendix B

Warfighter Interface Support

711 HPW/RH and Warfighter Interface Support

RH Network Support

Provided Network and Information Technology (IT) support to the Directorate.

1) General automation support for the following RH (Building 441), RHF (Building 441), IR (Building 29), CLN (Building 441), XP (Building 29), RHA (Building 441), RHX (Buildings 248 & 441).
2) Performed and monitored backup on 30 servers throughout the Directorate.
3) Continuation of Smart Force administrator training to satisfy AFI requirements.
4) Continued to push all patches to desktops that are missing patches. (All Branches)
5) Continued replacement and moving/consolidating data involving 6 servers.
6) Installed New Windows 2003 File Server for Bldg. 441, Transferred all Data from old server.
7) Reconfigured Bldg. 441 Old server to support internal Lab network.
8) Monitored past and upcoming patches to ensure compliance
9) Re-ghost every desktop and laptop in RH to a pristine AF SDC image to ensure software compliance.
10) Setup user mailboxes on 7 Canon copier/printers to disable unattended printing and begin printer consolidation effort
11) Provided support for several VTC/conference room sessions
12) Manually updated several desktop/laptop computers that failed TCNO checks; systems brought into compliance and migrated to SDC version 1.2; RHC now 100 percent SDC 1.2 compliant
13) Continued to roll out new computer systems- approximately 95 percent assigned new computers have been delivered
14) Continued to configure the systems for the core infrastructure of the RHC scientific network
15) Finished rollout of approximately 75 percent of new computer systems
16) Completed the building 146 refurnishing task, including:
17) Rewiring of the new cubicle areas in room 122
18) Moved the existing computer systems from building 190 back to 146
19) Set-up and checking out the computers after the move
20) Installed new Canon printers and removal of old printers
21) Set-up of user mailboxes and address book on new Canon printers
22) Repaired a Gateway LTO3 tape drive and put back into service
23) Prepared several systems / disk drives for turn in
24) Provided support for various conference room meetings
25) Provided desktop and printer support for the various RHC facilities
26) Continued to push all patches to desktops that are missing patches. (All Branches)
27) Fielded user queries
28) Replaced user computers
29) Upgrades to support Windows VISTA
30) Upgraded computers to make older computers available for turn-in
31) ADPE assistance
32) Supported SIPRNET users
33) Password changes, email setup
34) Supported Blackberry users
35) activation, password problems, account requests
36) testing Vista OS/Office 2007

37) KMC phone support
38) Delivered new computers
39) Supported SIPRNET users
40) Supported Blackberry users
41) CAC login problems due to client software online training
42) Updated software to allow users to complete
43) Increased number of users that required access to LiveLink for ERM software installs, CAC setup
44) CAC certificates
45) Problems with encrypted messages, digitally signing forms
46) New CACs, republishing certificates
47) Office/Outlook 2007 support, after the base push
48) Network account validations

RHC Computer Support

1) Performed backups and monitoring of 3 servers (700+GB of content) and 220 desktop systems
2) Provided general support to division users
3) Continued information collecting to support the construction of several certification and accreditation documents for RHC
4) Continued to assemble and configure the systems for the core infrastructure of the RHC scientific network
5) Installed 14 new network lines
6) Relocated 11 computers to the newly remodeled area of the 2nd floor, building 248
7) Relocated 2 printers to the newly remodeled area of the 2nd floor, building 248
8) Setup user mailboxes on 7 Canon copier/printers to disable unattended printing and began printer consolidation effort
9) Provided support for several VTC/conference room sessions
10) Manually updated several desktop/laptop computers that failed TCNO checks; systems brought into compliance and migrated to SDC version 1.2; RHC now 100 percent SDC 1.2 compliant
11) Continued to roll out new computer systems- approximately 95 percent assigned new computers have been delivered
12) Continued to configure the systems for the core infrastructure of the RHC scientific network

PROVIDE GRAPHIC SUPPORT FOR THE DIVISION CHIEF AND STAFF

1) Produced business cards for various government personnel
2) Continued producing new name plates for RHC
3) Shot, enhanced, and re-touched photos of RHC personnel
4) Designed new threatcon signage
5) Produced farewell montages
6) Designed and produced labels for CDs
7) Produced 30x40 posters for AtCat Lab Demonstration
8) Designed and modified graphics for DVED Imaging
9) Produced name badges for TTCP conference
10) Produced various signage for conference rooms
11) Completed installation of wall mural
12) Modified and produced new awards posters for lobby displays
13) Produced CD labels for TTCP conference materials
14) Designed and produced posters and lab signs for DICE Laboratory
15) Produced updated version of Awards Posters for lobby displays
16) Produced 3-D model of Battlespace for Gen. Bowlds' heraldic device
17) Produced signage for floor diagrams
18) Produced montage of RHC technologies for presentation

19) Collected graphics for RHCV wall poster
20) Completed HMD History for 8' wall poster
21) Modified graphic of C-130 Covert Landing for calendar
22) Completed graphics for framed, hall posters to represent RHCV technologies
23) Designed SAB icons
24) Shot, enhanced, and re-touched photos of RHC personnel
25) Updated building signage

RHCV General Labor

1) Preliminary design for Night Vision Logo (eagle)
2) Produced illustrations for transitional visor
3) Modified eagle log
4) Produced 3-D futuristic control station environment

Battlespace with Acoustic Support

1) Designed and produced illustration for Net-centric Audio
2) Designed layout for Spatial-Audio Display

BAO-BATMAN

1) Designed and produced BATMAN poster

CRISTL

1) Designed and produced banner for NASIC

RHCS Graphics Support

1) Designed and produced presentation slides

NETCentric Audio

1) Procured materials for display room in Building 441

FINANCIAL MANAGEMENT SUPPORT: Provide Financial Management support to the Warfighter Interface Division of AFRL

1) Cleared up errors, ULO, NULO and dormant records. Closed a total of 13 errors, ULO, NULO and dormant records totaling $24,349.01 during April 06
2) Interfaced with DCMA, Contracting and DFAS personnel to resolve discrepancies in the RHC accounting records
3) Responded to the specific financial tasking of the RHC financial management
4) Continued to produce customized financial reports using Cris, Mocas and Info Center to meet RHC financial management requires for specific and recurring financial data
5) Updated the RHC Civilian Payroll data upon receipt of the bi-weekly payroll data and reconcile it with the payroll date in the official DFAS accounting records
6) Updated daily the data capturing Small Business Innovation Research (SBIR) contracts belonging to RHC. It depicts the status of each contract obligation and expenditure as shown in the MOCAS and BQ accounting systems
7) Updated daily the RHC travel database maintained in Access
8) Kept RHC management informed of policy and procedures changes in the accounting/budget and help with the "what if" budget drills

GEN 4 Acces Plug Study

1) Two ad hoc subjects were scheduled and paid in support of the Gen 4 Acces Plug study
2) Ten subjects were scheduled and run in support of several attenuation studies conducted in the REAT facility, using Gen 4 Vented Acces earplug, BOSE earcups (active and passive ANR tests) and a 55-P helmet
3) Ear molds were scheduled and made for five ad hoc subjects
4) Data collection completed for the Gen 4 Acces Plug study

BOSE Study

1) Ten subject panel members were scheduled to run one REAT and one MIRE session in the BOSE study
2) Data collection for the BOSE study has been completed

3D Audio Chamber Studies

1) Subject panel availability and overall operation was monitored for the following studies:
   CRM Studies which measure the intelligibility for two types of synthetic CRM phrases in the presence of noise or other interferes.
   New Shinn Tail Noise Studies which assess the contribution of the reverberant portion of the target signal in identifying a target presented in the midst of multiple maskers.
   Grouping Studies which address questions about the relative salience of several cues such as on-set, fundamental frequency, common modulations and special location, and target segregation in multi-talker listening tasks.
   A Switch Studies which evaluate the relative importance of pitch of a target vs. ear of presentation in identifying a target in the presence of a speech masker.
   HRTF Studies which assess the efficacy of synthetically generated auditory horizon cues that will be used as an auditory display in GA aircrafts.
   Environment Test Studies which determines if there are meaningful words that can be used as warning signals, instead of sounds.
   MRT Angle Testing Studies which evaluate the extent of visual contribution (speech reading) in a speech intelligibility task as a function of viewing angle.
   Scaling Studies which evaluate the influence of a priori knowledge about the characteristics or content of the maskers or the target signal on a listener's ability to extract information from the target speech signal.
   Gun Exp studies which evaluate the effectiveness of a transparent hearing protection device by requiring the subjects to localize and identify a target phrase in the presence of gun fire.
   Cueing studies which evaluate the ability of listeners to detect and localize a target phrase which could be one of the following: forward PB words, reverse PB words, forward environmental sounds and reverse environmental sounds. The effectiveness of cueing will also be assessed by the presentation of a pre-cue or a post-cue.
   Tanya studies which assess the identification performance of listeners in the presence of two maskers which are 1) normal speech maskers, 2) Fo maskers and 3) Sineband maskers.
   Third Talker studies which evaluate the effect of a similar versus a non similar masker on target intelligibility.
   CRM_Detect studies which evaluate if detection thresholds differ as a function of the tasks that the listeners were require to do (for example, detect the presence of a target versus detect if the target is forward or reversed).
   Control_Dicho Detect studies which assess detection thresholds for a wide variety of tasks tested in CRM detect with and without a contralateral masker, and as the nature of the contralateral masker varies.
   Eavesdrop studies which explore the listeners' ability to detect call-back errors with two dyads (4 talkers) in a spatialized versus non-spatialized listening condition.
   Bands studies which evaluate the psychometric functions for two kinds of target signals: normal speech and filtered speech, in the presence of two other similar maskers.
   Detect Tone and Noise studies which validate thresholds.
   Whisper which evaluate target intelligibility with multiple whispering talkers, in order to assess target segregation efficacy in situations where takers are required to be unobtrusive.
   Bands_grouping which assesses the ability of listeners to identify a target signal under 3 experimental conditions: 1) when the target and masker had unique fundamental frequencies, 2) when the target and masker shared the same fundamental frequency, and 3) when the target contained some of the fundamental frequency information of the masker and vice versa.
   Grouping_control which assesses if the presence of a call sign aided target identification with artificial speech signals, where segregation was found to be difficult.

SpeedCP which assesses the influence of rate of speech on target segregation in a multitalker listening task.

CREARE Bone Conduction Study

1) Three subject panel members and five ad hoc subjects were scheduled for participation in the CREARE Bone Conduction Study
2) Data collection has been completed for the CREARE study

Marine NACRE Earplug Study

1) Two subject panel members and ten ad hoc subjects have been scheduled for the REAT portion of the NACRE study
2) Ten subject panel members have been scheduled for the ALF portion of the NACRE study
3) Data collection has been completed for 20 subjects (ad hoc and subject panel members) for the REAT portion of the NACRE study. Data collection is underway for the ALF portion of the NACRE study

UCAV Study

1) Four subject panel member and two ad hoc subjects scheduled for orientation, training and data collection.

56-P Helmet Study

1) Nine subject panel members participated in a study in the REAT facility in which the attenuation of the 56-P helmet was measured
2) Data collection has been completed

Combat Search and Rescue (CSAR) Study

1) Six subject panel members were scheduled in support of the Combat Search and Rescue study conducted in the CAVE facility
2) Data collection has been completed

ALF Localization 2

1) Ten subject panel members scheduled to participate in the ALF Localization 2 study; study completed.

Bandslocalization b

1) Ten subject panel members scheduled to participate in the Bandslocalization b study; study completed.

Bandslocalization 3b

1) Ten subject panel members scheduled to participate in the Bandslocalizaiton 3b study; study completed.

Fitts Study

1) Six subjects scheduled for the Fitts study; due to modifications in the design of the study, six subject panel members were scheduled to begin re-training for the Fitts study.

Counter Propaganda Study

1) Six panel members scheduled or voice recordings that will be used as stimuli for the MRT angle testing studies
2) All subject panel members scheduled for hearing tests
3) The new subject panel member was scheduled or earmolds for Gen 4 Asses earplugs
4) Weekly and monthly reports for tracking the amount of money paid to ad hoc subjects and panel subjects (cash payment during the probation period prior to hire) were prepared
5) One new male subject panel member was recruited and is working on a probationary period while his paperwork is processed

6) Quarterly security training, Privacy Act training, Records Management training and Information Assurance training was competed
7) All subject panel members received hearing tests

HGU-56P with Sound Guard Earplugs Study

1) Three subject panel members and seven ad hoc subjects participated in a study in the REAT facility in which the attenuation of Sound Guard Earplugs worn with a 56-P helmet was measured. Data collection has been completed for ten of twenty subjects.

Gen 4 Acces ANR with 55P helmet Study

1) Three subject panel members and two ad hoc subjects were scheduled to participate in this study in the REAT facility in which attenuation of Gen 4 Acces ANR with 55P helmet was studied
2) Data collection has been completed for five of ten subjects. This study is on hold since the earcups had to be returned to the company

Cueing_Exp

1) Eight subject panel members were scheduled to participate in 30 of 56 blocks of the cueing_exp study in the ALF Localization facility

Role of Real Time Auditory Feedback in a Delayed Virtual Environment Study

1) Four subject panel members were scheduled for the Role of Real Time Auditory Feedback in a Delayed Virtual Environment Study
2) Data collection completed

1279 Silynx Earplugs

1) Data collection completed for twenty subject

1280 Gentex HGU-56P with David Clark ANR Earcups with Acces Gen4 Aircrew Plugs

1) Five subject panel members scheduled to participate in a study in the REAT facility in which the attenuation of Gentex HGU-56P with David Clark ANR earcups was measured
2) Data collection complete

1233 Gentex HGU-56P with David Clark ANR earcups minus Custom Plugs

1) Seven subject panel members scheduled to participate in a study in the MIRE facility in which the attenuation of Gentex HGU-56P with David Clark ANR earcups was measured without the Acces Gen 4 aircrew earplugs
2) Data collection complete

Active Extreme Evaluation

1) Nine subject panel members and one ad hoc subject were scheduled to participate in a study in the REAT and MIRE facilities in which the attenuation of Active Extreme Earplugs was measured
2) Data collection complete

Adaptive Technologies (ATI) Earmolds

1) Seven panel subjects and one ad hoc subject had earmolds made in support of the ATI study

Joint Strike Fighter (JSF) Microphone Evaluation

1) Seven panel subjects were scheduled to participate in the JSF microphone evaluation in the VOCRES facility
2) Evaluation completed

Voice Recordings

1) Video and voice recordings under a whispering condition were scheduled for six subject panel members for use as stimuli material in the 3-D audio chambers

2) Data collection completed

Eyelink Study
1) Two subjects were recruited, scheduled and paid to participate in one session of the Eyelink Study
2) Data collection completed

AFRL/RHCB SUPPORT

Work includes several tasks for the Battlespace Acoustics Branch. Performing assigned duties in development and testing and fielding of acoustic protective and enhancement equipment. Working on programs utilizing Active noise reduction and cancellation. Developing and fielding the ACCES earplug system for the war fighter. Audio models data collection program is an on-going process in collecting Air Force aircraft noise levels. Assisting several program managers in on-going research studies and development work in audio acoustics. Appointed and performing duties as Branch PMEL Monitor, and Equipment Custodian for accountable and non-accountable equipment.

ACCES Program
1) Inspected several new ACCES cables for serviceability
2) Tested new ACCES stock cable to ensure quality of the product
3) Traveled to Whiteman AFB, MO, and collect noise data of the B-2 Bomber Aircraft. Data will be used in hearing protection procurement/design and the noise modeling program
4) Researched ANR headsets to be used in the VOCRES Lab
5) Modified several headsets for ACCES compatibility
6) Built 20 ACCES adapter cables for Seymour Johnson AFB personnel
7) Built 15 adapter cables for Capt. Divers at ACC Headquarters
8) Fabricated several new ACCES plus cables for several Generals and other DVs, personnel at Seymour-Johnson AFB, NC, Nellis AFB, NV, and Langley AFB, VA
9) Built two new microphone testing assemblies for the MIRE Facility
10) Traveled to Langley AFB, VA, and collected in-flight noise data on the F-22 and F-15 aircraft
11) Modified several 55P helmets and David Clark and Bose headsets for several Generals and other DVs
12) Collected noise data at the Wind Tunnel for modeling purposes
13) Modified a Bose ANR headset for connection to ACCES; gave headset and procedures to Bose for future duplication and creation of a modification kit
14) Ordered and received David Clark modification kits to modify Branch headsets
15) Re-vamped the ACCES Ear Plug headset attachment process and sent the new instructions out to required agencies

Dynamic Acoustic Models
1) Ordered and obtained material to build a new Microphone Calibration Speaker Assembly in the new anechoic chamber facility
2) Disposed of old equipment and material no longer needed
3) Took PMEL equipment to PMEL, calibrated some equipment in-house
4) Ordered equipment for new analysis and computer system
5) Scanned old data files to PDF files; DAT tapes, minidisks and reel to reel tapes transferred to wave files; this is being done in an effort to organize and catalogue past experiment in order to create a reference library
6) Tracked, received, and processed data collected; continuing on updating of equipment used on data collection process
7) Research and ordered new data collection media and other pertinent equipment

8) Finished sound pressure foam installation in the new Calibration Speaker System cabinet to be used by the Branch in future microphone calibration procedures/processes. After a thorough system test the speaker cabinet system will be installed in the ALF Chamber.
9) Collected aircraft cockpit noise data to be used in this program
10) A new speaker enclosure with a post was built to enhance microphone calibration procedures
11) Used Matlab to create Internal and External Microphone A-weighted Time Histories for the ten F-16 flights recorded at Cannon Air Force Base in early March 2006
12) Analyzed Time Histories and the In-Flight Acoustic Signature Data of each flight
13) Listened to each recording and noted all major events of the F-16 flights
14) Combined Time History data with the notes taken from listening to create and Excel spreadsheet for each flight documenting the "IN" and OUT" times for each maneuver or event
15) Completed data reduction for all ten F-16 flights at Cannon Air Force Base
16) Completed data reduction for all three T-38 flights
17) Data reduction for the B-2 flights is still in progress
18) Tracked, received and processed data collected. Continuing on updating of equipment used on data collection process
19) Researched and ordered new data collection media and other pertinent equipment
20) Assisted AFIT personnel in recording the SARL wind tunnel for noise reduction research
21) Assisted in the set-up and calibration of the TEAC recording system
22) General housekeeping duties (i.e. sweeping, dusting, removing unused boxes and other unused items from area)
23) Arranged and organized workstations
24) Research and ordered more equipment for new analysis computer
25) Assisted in analyzing data from JSF tests
26) Completed itemized listing for Bob McKinley
27) Create a digital text archive for Air Force Noise Measurements Data by scanning all related reports, graphs, charts, data, protocols, pictures, etc.
28) Create a digital audio archive by converting all DAT tapes, Mini Discs, 16 channel TEAC tapes, and reel to reel tapes associated with the Measurements Data to .Wav files
29) Assisted in analyzing data from JSF tests

Db Towers

1) Attended several meetings on possible design and function of the db Towers system
2) Researched steel purchase and started process of ordering required equipment
3) Went TDY to Paducah KY to attend a meeting with World Tower Inc., company that is subcontracted to install the Towers for us
4) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project
5) Several meetings with B & K Inc. and research time spent looking for a compatible Audio accusation system
6) Several more meetings on Towers design and system layout
7) Currently working funding issue / purchase order for steel storage
8) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project
9) Received quotation reports on 300 foot and 1000 foot tower installations
10) Attended several meetings with National Instruments and Audio System Training sessions with B & K Inc. personnel. Research time also spent looking for a compatible Audio accusation system
11) Attended several more meetings on Towers design and system layouts
12) Still currently working funding issue / purchase order for steel storage
13) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project
14) Received quotation reports on 300 foot and 1000 foot tower installations
15) Attended several meetings with National Instruments and Audio System Training sessions with B & K Inc. personnel. Research time also spent looking for a compatible Audio accusation system

16) Attended several more meetings on Towers design and system layouts
17) Still currently working funding issue / purchase order for steel storage
18) TDY for 5 days to White Sands New Mexico for Tower Conference and Noise Data Collection and testing at proposed project site
19) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests completed for material and support required
20) Received more modified quotes on the 300 foot and 1200 foot tower installations and temporary steel and accessories storage
21) Attended several more meetings with National Instruments and Audio System Training. Continued research time also spent looking for a compatible Audio accusation system
22) Attended several more meetings on Towers design and system layouts
23) Currently working Environmental Assessment, Archeologist, and Survey quotes
24) Audio models data collection for Air Force Aircraft noise levels.
25) Assisting with Concept Operations Plan for towers project
26) Built a trolley system with rails and a cart for the 300 foot tower project
27) Bought and sent metal to World Towers, Inc. for testing and proof of Concept
28) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests completed for material and support required
29) Received more modified quotes on the 300 foot and 1200 foot tower installations and temporary steel and accessories storage
30) Attended several more meetings and telephone conferences on Tower construction and proof of concepts
31) Ordered several systems for continued testing of audio collection systems
32) Attended several more meetings on Towers design and system layouts
33) Still working Environmental Assessment, Archeologist, and Survey quotes
34) Assisting with Concept Operations Plan for towers project
35) Built a trolley system with rails and a cart for the 300 foot tower project
36) Bought and sent metal to World Towers, Inc. for testing and proof of Concept
37) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests completed for material and support required
38) Received more modified quotes on the 300 foot and 1200 foot tower installations and temporary steel and accessories storage
39) Attended several more meetings and telephone conferences on Tower construction and proof of concepts
40) Ordered several systems for continued testing of audio collection systems
41) Attended several more meetings on Towers design and system layouts
42) Still working Environmental Assessment, Archeologist, and Survey quotes
43) Assisted in the review and completion of the DOPAA and Concept of Operations for the tower installation
44) Designed a Trolley System for the towers and submitted it for design and work to a local machine shop for development
45) Researched and ordered National Instruments equipment, and microphone systems for potential use in the Tower project
46) Researched and ordered equipment for a test of the 300 foot tower trolley system
47) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests including the Towers themselves were completed for material and support
48) Ordered and received cable, pliers, piping, fittings and clamps for cable installation for the tower project
49) Worked new storage area for the steel and accessories for the 1200 foot tower
50) Attended several more working sessions National Instruments and Audio System Training, and continued research on a viable system for audio data collection for the towers project
51) Attended several more meetings on Towers design and system layouts
52) Still working with the Environmental Assessment, Archeologist, and Survey personnel at White Sands Missile Range

53) Currently working and ordering the last of the equipment for the 300 foot tower trolley system that I designed
54) Ordered and had delivered 2 off road vehicles and wagons for the towers projects
55) Researched and ordered equipment for a test of the 300 foot tower trolley system
56) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests including the Towers themselves were completed for material and support
57) Ordered and received cable, pliers, piping, fittings and clamps for cable installation for the tower project
58) Worked new storage area for the steel and accessories for the 1200 foot tower
59) Attended several more working sessions National Instruments and Audio System Training, and continued research on a viable system for audio data collection for the towers project
60) Attended several more meetings on Towers design and system layouts
61) Still working with the Environmental Assessment, Archeologist, and Survey personnel at White Sands Missile Range
62) Currently working and ordering the last of the equipment for the 300 foot tower trolley system that I designed
63) Ordered and had delivered 2 off road vehicles and wagons for the towers projects
64) Researched and ordered equipment for a test of the 300 foot tower trolley system
65) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests including the Towers themselves were completed for material and support
66) Ordered and received cable, pliers, piping, fittings and clamps for cable installation for the tower project
67) Worked new storage area for the steel and accessories for the 1200 foot tower
68) Attended several more working sessions National Instruments and Audio System Training, and continued research on a viable system for audio data collection for the towers project
69) Attended several more meetings on Towers design and system layouts
70) Still working with the Environmental Assessment, Archeologist, and Survey personnel at White Sands Missile Range
71) Currently working and ordering the last of the equipment for the 300 foot tower trolley system that I designed
72) Ordered and had delivered 2 off road vehicles and wagons for the towers projects
73) Researched wench system for the 300 foot tower trolley system
74) Completed the 25 foot tower / scaffolding in the basement of bldg. 441 to serve as a tower platform to test trolley system I designed and future microphone positions and applications
75) Ordered more material and supplies for this project
76) Attended several more working sessions for on viable system for audio data collection for the towers project
77) Attended several more meetings on Towers design and system layouts
78) Still working with the Environmental Assessment, Archeologist, Coring Company, and Survey personnel at White Sands Missile Range; on hold for heavy equipment rental
79) Processed several purchase requests for this project
80) Researched wench system for the 300 foot tower trolley system
81) Completed the 25 foot tower / scaffolding in the basement of bldg. 441 to serve as a tower platform to test trolley system I designed and future microphone positions and applications
82) Had trolley system modified by Quality Machine Shop
83) Continued working on getting a new storage area for the steel and accessories for the 1200 foot tower, assisted in cleaning out area at bldg.64
84) Attended continuing training sessions National Instruments and Audio System Training, and continued research on a viable system for audio data collection for the towers project
85) Attended several more meetings on Towers design and system layouts
86) Assisted with the Environmental Assessment, Archeologist, Coring Company, and Survey personnel at White Sands Missile Range, had Assessment copies printed
87) Processed several purchase requests for this project
88) Tested cable and connectors for data collection system

89) Traveled to Socorro, New Mexico to survey the future site of the dB Towers at White Sands Missile Range.
90) Placed markers to frame the control center building and parking lot.
91) Marked the road going from the access road to the control center.
92) Marked the location of the Towers and the center of the recording array.
93) Marked all other microphone locations along the array.
94) Picked up trollies that were modified by Quality Machine Shop
95) Assisted in setup and marshaled 3 Loads of steel from the Kentucky plant to our new storage area, bldg.64
96) Attended several planning meetings for the ARC Complex and research on a viable system for audio data collection for the towers project from the National Instruments Company
97) Attended several more meetings on Towers design and Power grid system layouts
98) Designed Power setup for entire ARC facility and submitted to Program Manager and potential installation contractor
99) Assisted with the Environmental Assessment, Archeologist, Coring Company, and Survey personnel at White Sands Missile Range, had Assessment copies printed
100) Purchased material and started the fabrication process for microphone box systems

BAM Lab
1) Cleaned out old BAM/ARTD equipment from BAM Lab
2) Installed overhead projector rail system
3) Three subject panel members participated in a Speech versus Test Demo study. Data collection completed

BAO Program
1) Built/customized a Laser Range finder to work in tandem with the BAM Lab video screen/control software. Modification consisted of major joystick circuit board modifications, and installation of a new viewing system with USB and VGA video cable connections
2) Research and ordered radio equipment and GPS antennae for BAM Lab

Multi-source Sound Localization
1) Maintenance and upkeep of the Auditory Localization Facility (ALF) Chamber
2) Installed track guide wiring and stops and several track sections into the new race car track in the new anechoic

Audio Displays and Speech
1) Built several amplifiers from kits for use in on-going Audio and Speech research
2) Obtained material for modification of new anechoic chamber facility to support a new speaker rail system for future experiments
3) Assisted in making HRTF recordings (Head Related Transfer Functions)
4) Assisted in troubleshooting with the ALF (Audio Localization Facility) operating system
5) BandwidthStudy3, BandwidthStudy4 and BandwidthStudy5 are complete

Informational and Energetic Making
1) Traveled to Fort Knox, KY, and collected noise data recordings on the M240 and M249 Machine Guns to build a future noise data compatibility table
2) Set up talker/listener experiment in the VOCRES chamber
3) Installed several new serial cables in the Anechoic Chamber/ALF chamber for the wing speaker tests
4) Performed new tests on a new amplifier for the chamber amplifier rack
5) Ordered material to build a Subject Response LED assembly to speaker audio output drive connection box for upcoming experiments by Dr. Brungart
6) Completed four station VOCRES modification, installed internet wiring, monitors with BNC cable connections and new cameras
7) Wired new audio cable for VOCRES stations, performed modifications to console hardware
8) Re-vamped the Subject sitting seat for the ALF Anechoic Chamber

9) Built a set of Panasonic microphones for testing in the ALF chamber
10) Research and purchased a set of laser levels for the ALF Chamber
11) Built two cable for the RF Radio testing with the Min-Tac System
12) Removed old equipment from Lab and control room
13) Manufactured microphone set-ups and power boxes for ALF tests under this program
14) Modified microphone set-up for calibration purposes

Info-Energy Mask

1) Four VOCRES stations were upgraded with network cabling to the control desk. Each of those stations received six video cable connected among the other three stations

Branch Support

1) Researched and obtained required tools and equipment needed for Branch studies
2) Discarded old and outdated equipment throughout Building 441
3) Installed new monitor and camera in the REAT chamber
4) Obtained material and built a large display board for the MIRE and REAT chamber facilities
5) Completed purchase requests for material needed for facility/Branch area repairs
6) Took PMEL equipment to PMEL, calibrated some in-house
7) Continued cleaning out area down stairs to facilitate further enhancement of the BAO facilities and system program
8) Cabinets containing electronic components were sorted
9) Inventory of old computers begun for UCI inspection
10) Upgraded the MIRE Lab facility with a new G.R.A.S. microphone system
11) Created a High Value storage area for tools and other things of value
12) Built five Tool Kits for the Branch
13) Created an Electronics Lab for the Branch
14) Installed EMI equipment in the new Branch Electronics Lab
15) Completed preparation for QA inspections and cleared errors found by inspectors

Audio and Text Archive

1) Created a digital text archive for Air Force Noise Measurements data by scanning all related reports, graphs, charts, data, protocols, etc.
2) Created a digital audio archive by converting all DAT tapes, Mini Discs, and reel to reel tapes associated with the Noise Measurements to .Wav files
3) The complete contents of one of the four filing cabinets due to be converted was scanned
4) Scanned the complete contents of one and a half of the four filing cabinets due to be converted
5) Created a digital text archive for Air Force Noise Measurements data by scanning all related reports, graphs, charts, data, protocols, pictures, etc.
6) Created a digital audio archive by converting all DAT tapes, Mini Discs, 16 channel TEAC tapes, and reel to reel tapes associated with the Noise Measurements Data to .Wav files
7) Scanned the complete contents of three of the four filing cabinets due to be converted
8) Approximately 95 percent of the reel to reel tapes have been converted to .Wav files
9) Approximately 85 percent of the pictures and picture negatives have been converted to JPEG files.
10) Approximately 95 percent of the DAT tapes have been converted to .Wav files.
11) Approximately 75 percent of the Mini Discs have been converted to .Wav files.
12) Approximately 95 percent of the reel to reel tapes have been converted to .Wav files.
13) Approximately 55 percent of the 16 channel TEAC tapes have been converted to .Wav files
14) Approximately 25 percent of the DAT tape slip covers have been scanned as JPEG files

PC Based 3-D Audio Rendering

1) Changes were requested and accomplished for the Orientation HRTF Validation Study. A preview feature has been incorporated into the GUI so that the cue level can be adjusted for

the comfort of the listener and also controls to effect the various attitude changes so that the subject can hear what the various cues sound like. The sense of the pitch cue was reversed and the correct responses for all were adjusted so that the correct response is the action needed to be taken to arrest the indicated change in attitude (i.e. up for a pitch down; right for a roll left). The cue onsets were also made to start flat (wings level) and then to slew to the final target position within a two second window. The rate required for the largest displacement was calculated for the two second window and all other targets were slewed at the same rate for that plane. The host computer for this task again began exhibiting problems with audio streaming device that now need to be resolved before the study can commence. Also the PI will be looking into changes needed for the HRTF data sets to be validated to correct some lateralization issues in the pitch HRTFs.

2) Three new commands were added to the IPSS to support slewing or rotation of the head (listener) position in the non-head-tracked mode. The commands allow for the specification of a start and stop position in the one of three orientation (rotation) angles and the increment to be applied for every (approximately) 20 millisecond interval. The commands have been implemented in function for the Orientation Study requirement discussed above, but presently contain no error checking or error status return processing. This will be accomplished soon and a new version released with supporting documentation. A mute SLAB error display feature was added to an earlier version (for SLAB 5.1.4) of the Audio Server library. This allows for the suppressing of the missing file error message when the GA flight test programs (simulation and real) are cycling through orientation cue wave files; however, any errors are still recorded in the log file during the mute period. This feature will need to be migrated to the latest (SLAB 5.7.0) version.
3) The HRTF Validation follow-on study was concluded this period. A HRTF dataset for the attitude cue has been validated for use in the GA-based flight test to be conducted at NASA Langley beginning in April.
4) No additional support issues were required for the IPSS use in the Cave CSAR experiment and demo.
5) An interface module for the Polhemus 3-Space and IsoTrak head trackers is being developed for the (IPSS) Internet Protocol SLAB Server. In addition to the communication module, support for head position updates in Cartesian coordinates is also being worked. The final design will provide the capability to maintain head location in orientation angles only, and in Cartesian coordinates and orientation angles at the same time. This will maintain compatibility with existing systems that only require the orientation angle updates. To facilitate implementation, a new head tracker type has been defined for the IPSS configuration file. To select the Polhemus tracker COM port and baud rate, comma delimited values are supported in the configuration file head tracker type line. A final comma delimited parameter on the head tracker type line defines the head update mode (Cartesian with orientation angles or orientation angles only); the head update mode parameter can also be employed for the previously supported head tracker types. The default head position mode is orientation angles only. This effort is about 50 percent complete.
6) Support reinstalls of audio server (IPSS) and software dependencies (SLAB, DirectX, etc.) following rebuild of OS on the audio server PC in the Cave.
7) Modifications and enhancements for the (IPSS) Internet Protocol SLAB Server to support upcoming experiments in VOCRES are nearing completion. The Polhemus head tracker type interface is working and support for six DOF including Cartesian coordinates head position updates, has been verified and validated. Although designed to support the other tracker type interfaces also, six DOF position updates has not been tested for the InterSense and the TrackD API types. To allow for rotation of coordinates systems between the source and

sensor two configuration file parameter keywords have been defined, namely "orientation" and "xyz" which allow for the specification of a rotation (sign) vector for each of these coordinates; the default vector is (1, 1, 1) for each. An IPSS command has also been added that supports setting a boresight reference to preclude performing a boresight function, and provide an appropriate reference automatically upon client program initialization. A companion query command is also provided to return the boresight reference angles. The dependency for the TCP/IP interface library TCP4U has been removed and replaced with the Windows IP socket class. This new version (v2.2) of the IPSS also uses SLAB 5.8.0 (the current release version of SLAB). The recent (April 2006) release of DirectX has also shown to be more compatible with SLAB and no longer requires the installation of the patch provided by NASA Ames for earlier versions of DirectX. All new features and dependency requirements have been documented in the IPSS User Reference Manual, which is available in the shared development folder of the lab network server.

8) Installation of the new audio server (IPSS) and software dependencies SLAB 5.8.0 and DirectX (April 2006) on the audio server PC in the CAVE has been accomplished.
9) Changes made to the (IPSS) Internet Protocol SLAB Server in support of the VOCRES audio-video correlated study are complete and validated. An on-command capability has been added to the IPSS to support FIR tap length HRTF datasets with up to 256 coefficients. A copy has been distributed to the CAVE monitor for use in new and existing applications. All changes have been documented in the IPSS Users Reference and is distributed on the shared development folder of the Lab LAN server.
10) A new HRTF Validation Study (follow-on) was performed to evaluate and validate a 256 tap length attitude dataset for potential GA experiments. The new version of the IPSS was used to render the sound sources with the long tap length dataset. In support of this effort, the IPSS was installed on a new PC system designated for subject self-paced experiments; checkout with the new Audigy 5 sound interface was required. Initial failures were traced to an incorrect programming of the sample rate for the interface. Although the Audigy 5 defaults to a 44.1 kHz sample rate, the Audigy 5 will only work with SLAB at 48 kHz sample rate (the default sample rate for all other Audigy devices the lab has encountered). Once programmed for the 48 kHz sample rate, all checked out properly and the study was allowed to run.
11) Some support has been provided for the integration of the latest version IPSS in the CAVE. The audio noise associated with rapid movement of the rendered sound source continues to be apparent. The CAVE facility software developer was asked to investigate the phenomena using the SLAB utility SlabScape and manipulate the SLAB smoothing parameter to see if any effect can be observed (IPSS sets this parameter for no smoothing). The developer has reported that no difference was observed and that the noise is not apparent when SlabScape is employed. No further action will be taken without the direction of the government Task Monitor.
12) The 256 tap length HRTF validation study has been completed successfully.

Langley Flight Test

1) Attention was given to completion of the GA flight test simulation study to achieve the final (simulation) validation of the flight test software. Changes required have been in the display of the orientation (attitude) cue and the sense of the pitch cue. The attitude cue has been made to be rendered always with respect to the plane attitude and independent of the subject's head position. To effect this it was necessary to change the manner of presentation for both cues (attitude and direction) in the head-slaved operation mode. The cues are moved with respect to the listener who is left fixed at the (0, 0, 0) position. The control program uses the vector math class to transform the position of the sources as the plane

(attitude) or head and plane (direction) position change and reposition the sources at the transformed locations in real time. Other changes requested include the reversing of the sense of the pitch cues and some changes to the formatting and content of the text data files to support analysis. Orientation source files have been generated for individual subjects from sources provided by each subject. Several subjects have been run in the heading (HDG) task to date. Data collection will continue into the next period for all three tasks (CIA, RUA and HDG). Several telecoms with NASA have been participated in for flight test program planning and execution as required.

2) The Cave simulation resumed and revealed some further changes and corrections needed for the GA Flight Test control software. A Suspend and Resume feature has been added to the (HDG) Heading or Navigation task to allow for repeating a waypoint due to interruption from encroaching traffic or other condition. The tab order of the controls was cleaned-up and the controls made responsive to a keyboard space bar strike. Sound cue levels for the (CIA) Change in Attitude Threshold Detection and (RUA) Recovery from (unusual) Displaced Attitude tasks were adjusted for proper display and plane-slaved mode was made the default condition for those tasks (head-slaved condition will not be run in the experiment). The CIA command generation was corrected to provide a truer random selection of the commands, and to display the command prior to the start of the trial. However, after flying the initial ICF flight at NASA it has been decided to move these commands to the flight card for the task. This eliminates cumbersome switching of the intercom modes during flight.

3) Traveled to NASA Langley to conduct (ICF) Instrument Check Flights to validate the experiment control software and interfaces in the plane. Changes were made to the software following these flights to correct the sense of the roll cue in the head-slaved mode for the HDG task, clean up the flowing in and out of tasks without a program lock-up, and providing a means of selecting audio or no audio for the RUA task. Other issues revealed involved the calculation of head-slaved target positions in the clockwise coordinate system and then presenting in the SLAB counter-clockwise coordinate system was not being handled properly, and the display of the roll attitude cue was reversed in the head-slaved mode. Fixes have been coded; however, not all have been tested in flight. This will be accomplished during the upcoming TDY scheduled for the second week of April 2006. Issues yet to be resolved are the management of the CIA trials (audio and non-audio; with and without instruments), and the proximity tolerance value for the HDG task. Another unresolved issue is the dropping of head position data from the data stream provided by the PC104 head tracking system; the cause and resolution (if any) of this anomaly resides in the PC104 software program. The work-around for this is to stop and restart the PC104 software when this happens. Maintenance of the software including updates to the flight computer and support of the ICF conduct was provided as needed; one ICF conduct was participated in as an experimenter. Generation of additional sound wave files for the attitude cues for added subject numbers was begun and installed to the demonstration laptop; these will be installed to the flight computer on the next trip. Subject provided audio wave files will be generated as the source media are made available. Provided support for the Cave GA simulation experiment data collection and analysis as required.

4) Traveled to NASA Langley the week of 10 April 2006 to support the final (ICF) Instrument Check Flights and the start of data collection flights. Various trials to resolve the plane-slaved navigation cue issue were conducted in the hangar and on the tarmac by towing the plane with a tug while listening to the navigation cue. A version that seemed to checkout on the ground failed to perform correctly in-flight on two separate occasions; data collection in plane-slaved mode was precluded. The final resolution of the problem was to physically

mount the head sensor in the plane and conduct those flights with the head-slaved mode of the software. Modifications to the software were made to conduct the plane-slaved mode trials with head-slaved processing; one or more plane-slaved mode data flights were flown before the end of the week.

5) The proximity tolerance for the navigation task was increased to approximately 0.3 nm. Two ICF flights were participated in; participation in pre and post flight briefs as required were accomplished; and support for data collection flights was provided as needed. Software tracking by NASA was complied with for all software update releases made and installed in the flight computer during the week.
6) A failure of the head tracker system required the removal from the pane and return to WPAFB. Working in conjunction with AFIT/ENG, it was determined that a BIOS setting defining the boot device had become corrupted. Resetting the parameter abated the failure and the head tracker system was returned to NASA by overnight courier so that it might be installed and checked out by NASA in readiness for data collection flights scheduled for the week of 01 May 2006.
7) An apparent saturation of the file structure of the data collection folder on the flight PC (GPC #3) resulted in run errors in the experiment control program. A suggestion was provided to the experimenter to archive and remove all of the current data from the folder and to run the program with a clean data folder. This suggestion proved to provide a work-around for the run errors. The data collection flights have been concluded at NASA Langley for the current phase of the study; however, some make-up flights may be scheduled from an alternate airport location.

Soft Phone Data Stream Rendering

1) The lab machine that is installed with the sipXpbx soft phone system is currently displaced from its former location because of restructuring and reallocation of lab work areas. Work on the soft phone system is suspended until the restructuring and associated moving is completed.
2) A room has been designated by HECB to become the networking lab. Among other jobs to be supported will be the SIP PBX capability. The lab machines currently designated as the SIP PBX server and a client machine have been moved to this room. Continuance of work is pending the setup of the networking lab to be accomplished.
3) A request notice to the HECB Lab network monitor for the DNS configuration requirements in the Lab server to support SIP PBX has not been responded to. The same request will be forwarded to the local CSA in hopes that appropriate action or course of action can be obtained.
4) A review of an open source Voice over IP library (JVOIPLIB) is in progress as a possible replacement for the SIP approach which is presently stymied over server requirements. As directed, work to install and exercise the library in the networking lab may be planned.
5) Several PCs slated for possible turn-in have been picked and assigned to the networking lab to add to the existing resources already there. This brings the total of resources to five PCs; another older PC is being retained for historical/archival purposes as it supports the KEMAR method of processing raw HRTF data.
6) A Dell laptop computer previously used in support of GA flight test software development was cleaned up and turned over to an extra-branch developer for VoIP transmit software development. The originally planned laptop for this effort was found to have been returned from a secure environment missing a hard drive and mounting bracket. A search was conducted to determine the availability and cost of a new hard drive and hardware. Rather than replacing the hard drive at this time, it was decided to identify an alternate laptop; the selected laptop is the Dell Inspiron 5100. All pertinent software projects were archived and

other files no longer needed were removed to free up space for VoIP development. The laptop computer has been signed out to the developer for a period of one year.

7) The new vu-meter version of the Spatial render plug-in for SLAB and a corresponding demo application has been delivered by VRSonic.

VOCRES Audio Experiment

1) The Gateway 600 laptop computer was used in the VOCRES chamber to test and verify the operation of the IPSS rendering three sound sources in Cartesian coordinates (six DOF). Head tracker updates in Cartesian and orientation coordinate systems were provided by a Polhemus 3Space tracker. Upgrades and corrections for the IPSS needed to support this requirement were explored and tested and verified. The laptop platform hosted both the IPSS and a test utility audio client application modified to control the presentation of the three sound sources correlated with the location of three video monitors mounted on the subject station. The client application was also used to map the location of these video monitors by mounting the head tracker sensor on a headset and placing same on or near the monitors and querying the IPSS for head position data. The configuration of the test audio client application will be used to model development of a specific audio client application for the conduct of this phase of the experiment.
2) Development of an Audio Client application for the VOCRES experiment is in progress. Ultimately it will control spatial rendering of up to six audio sources and interface to the IPSS running on the same platform. Immediately, the IPSS will be made to present three sound sources in Cartesian coordinates that correlate with the physical location of the three video monitors with respect to the head tracker source mounted at the subject station. The sound sources will be head-coupled to the listener with six DOF head position updates rendered by the IPSS. To this end, and in anticipation of the not yet available SoundBlaster Audigy 4 audio interfaces, the ASIO4ALL WDM audio interface has been employed to support development and testing of at least two of the sound sources. As soon as a dedicated subject station PC with the Audigy 4 becomes available, development will proceed with the target hardware. This work is about 50 percent complete.
3) Four PC systems have been identified and configured to use for the study. The VOCRES client application has been completed and has been made to spawn the IPSS and delay for an appropriate amount of time to achieve a connection; the systems are configured to automatically perform this task after log-in. The application then begins streaming of the, currently, three audio channels via the ASIO driver support of the Audigy 4 sound interface; the sound channels are spatially located to correlate with the position of the three video monitors. Head tracking is achieved via the Polhemus interface built in to the IPSS. The VOCRES client application and the current version of the IPSS have been installed on all four systems. Each is configured with the Audigy 4 sound interface for audio streaming of the three voice channels. Each system is also configured for remote desktop so that all systems may be controlled from a single keyboard, mouse and video monitor utilizing the remote desktop connection. Full system checkout (excepting head tracking) has been accomplished using the Networking Lab as a staging area. All four systems will be relocated to VOCRES when the appropriate audio cabling has been installed there.
4) The four designated and configured audio computers for the VOCRES A/V experiment have been relocated to the VOCRES control room. The newly installed audio cables and head tracker serial cables from each desk have been routed to the appropriate audio PC. Still lacking are a hub and/or LAN lines to network the PCs and the input audio cables to run from the station intercom panels. This work should be completed early in the next report period.
5) The four audio PCs are now networked and audio line feeds from the subject stations are now run and connected to the individual Audigy audio interfaces as required. Check out of

the operation of the processed streamed audio was performed. To enhance the checkout experience the software was temporarily modified to display the audio channels at exaggerated left and right positions so the placement could easily be discerned without the benefit of head tracking, which is still lacking from the setup. The head tracking capability is pending sensor mounted headsets availability. The streaming of the appropriate audio channels at each station and the rendered positioning of each audio channel has all been verified.

6) The JVOIPLIB components were downloaded and installed on the development environment PC. It was determined that the components required the 2005 version (v8.0) of the MS Visual Studio C++. As no 2005 license is available at this time VS Express 2005 versions were downloaded from the MS Developer's Network (MSDN). The Express versions proved suitable for compiling the library components but were unable to compile and build the test utility, an MFC application. Attempts to convert the application to a 2003 development environment have proven unsuccessful. Further evaluation work with the JVOIPLIB is pending the availability of, or access to, a 2005 MSVS license.
7) The checkout of the four audio streaming data channels to the four subject stations, with head tracking, was successfully accomplished and demonstrated this period.
8) Concurrent with the verification of the VOCRES audio/visual head tracking function, the 6-DOF head position updates has been shown to work using the InterSense API.
9) Support for a VOCRES demonstration to visitors was performed during this period; required report documentation was accomplished.

Acoustic Signal Control: Program utilizing active noise reduction and cancellation

1) Modified several David Clark and ACOUSTICOM Headsets for connection to ACCES, gave headset and procedures to David Clark and ACOUSTICOM for future duplication and creation of a modification kit
2) Built a ear plug system complete with power and filter applications for upcoming testing to be completed on the new and current ACCES configuration
3) Demoed modification on video for training of other Air Force Base Life Support personnel in completion of headset modification
4) Assisted with Concept Operations Plan for towers project
5) Built a trolley system with rails and a cart for the 300 foot tower project
6) Bought and sent metal to World Towers, Inc. for testing and proof of Concept
7) Constant telephone and email contact with lead engineer and salesman on completion of Tower Project. Several purchase requests completed for material and support required
8) Received more modified quotes on the 300 foot and 1200 foot tower installations and temporary steel and accessories storage
9) Attended meetings and telephone conferences on Tower construction and proof of concepts
10) Ordered several systems for continued testing of audio collection systems
11) Still working Environmental Assessment, Archeologist, and Survey quotes
12) Met with Vernie Fisher on the basics of operating the MIRE facility.
13) Shaped a "pink noise" sound wave to specification for the NASA space suit study.
14) Changed blown speakers and replaced blown fuses in the MIRE facility.
15) Helped with the preliminary setup for the device Helmet Localization study in the ALF facility.
16) Attended more meetings and discussed new engineering and procurement techniques to bring about the new Active Noise Reduction ANR Custom fit communications plug system to the USAF aircraft cockpit
17) Completed ACCES to David Clark and Bose communication headset modification procedures for Air Combat Command Life Support Personnel
18) Attended meetings and discussed new engineering and procurement techniques to bring about the new Active Noise Reduction ANR Custom fit communications plug system to the USAF aircraft cockpit
19) Ordered several different more sets of ACCES plugs for testing and Individual issue
20) Modified several helmets and ordered and modified several headsets for C-17 pilots at Charleston, AFB SC and Scott AFB, IL

21) Ordered several more sets of ACCES plugs for various Generals, Colonels and other personnel
22) Ordered supplies to modify helmet an headsets of pilots of C-17 aircraft at Charleston AFB, SC
23) Learned to calibrate and operate the REAT and MIRE facilities
24) Ran subjects in the REAT facility for the JSAM chemical defense mask and helmet study

3D Global Hawk Communication Study

Project Status Summary: The test objective is to measure the effects of continuous variable slope delta (CVSD) and in tandem with CVSD, adaptive differential pulse code modulation (ADPCM), and voice over internet protocol (VoIP) vocoding algorithms on speech intelligibility over an ARC-210 radio link. These components are considered the critical links in the air traffic controller (ATC) to global hawk mission control element (MCE) communication path. Generally, the guidelines in ANSI S3.2-1989, measuring the intelligibility of voice communication systems, will be followed by using the modified rhyme test (MRT). Speech intelligibility will be evaluated with the ARC-210 radios in non-secure, secure, and HAVE QUICK II modes in the simulated communications path between the ATC and MCE stations and in real communications via an INMARSAT communications link. The communication system must achieve a required mean intelligibility level of 80 percent with the MRT to be considered acceptable by the operator.

Intelligibility Measurements of CVSD, ADPCM and VoIP Vocoding Techniques

1) Work with the AG2000BRI card has been suspended in favor of a TI DSP platform utilizing vocoder algorithms written for the TI DSP processors. A search of available in-house DSP resources has located one TI C5510 DSK card and one TI C6711 DSK card. There are also known TI C6211 DSK cards installed with 3-DVALS units (two with each). The C6211 cards associated with the 3-DVALS development board as well as the acquired C6416 card have been discarded during the dismantling of the former Electronics Lab. As noted the platforms which have the most utility with algorithm vendors are the C55xx and C64xx families. Adaptive Digital Technologies (ADT) has been chosen as a vendor to provide evaluation licenses for an ADPCM algorithm and a MELP algorithm. Until the algorithm libraries can be purchased and delivered, interim work is being accomplished to adapt source code for a MELP algorithm, available from the (DDVPC) DoD Digital Voice Processor Consortium web site. The algorithm was originally written for C5x, however C source code is available. The initial step was to implement, using TI DSP/BIOS, a two stage (encoder and decoder) audio pass-through application. Once this was working, work of porting to the C5510 platform using TI DSP/BIOS was started. Presently the two stage MELP algorithm will build successfully and load, but is not yet executing properly. This work, even if not successful, will serve as a basis for implementing the ADT algorithms, when available.
2) An attempt was made to build the JVOIPLIB libraries and test utility to evaluate the performance of JVOIP. The libraries and test utility require MS VS C++ 2005 (version 8). Although the libraries would build with the freely available VS 2005 Express, the test utility, an MFC application, would not. Access to a computer with an available license for the full version of VS 2005 is currently being arranged.
3) Prototype applications for performing ADPCM encoding and decoding of analog signals sampled at the line input of the C5510 board were successfully implemented this period. Two channels of data are being processed on the board. Sample rates of 16 kHz, and 32 kHz are being implemented in separate demonstration applications. The same application framework implemented for the MELP applications is being employed in the ADPCM applications. This framework allows for control of the sample rate and the processing of the interleaved audio data from the two channel (ADC) analog to digital and (DAC) digital to analog converters.
4) A-law and u-law companding algorithms were developed for the C5510 from a TI application note white paper describing an assembly language implementation for a C54x environment. The routines are coded in C and were debugged and validated using test data and equivalent

codes provided by the application paper. Both companding methods are available; however, presently the selection of the companding method is accomplished by an edit, compile and build sequence of the prototype application. Likewise the selection of the desired sample rate also requires an application generation cycle.

5) Other work this period involved the development of up sampling and down sampling algorithms. In order to generate a 16 kHz signal, down sampling and subsequent up sampling needed to be employed as the ADC and DAC converters do not directly support a 16 kHz sample rate. In conjunction with the re-sampling algorithms it was necessary to employ a low pass filter to reduce aliasing of the sampled data. The Atlanta Signal Processors (DFDP) Digital Filter Design Package was utilized to design a (KFIR) Kaiser Window Finite Impulse Response low pass filter with the appropriate pass band and stop band cutoff frequencies. Several filters of different tap lengths were generated; the longer tap lengths provide tighter control on the ripple effect at the cutoff frequencies. The resultant filter coefficients are input to TI DSP library filter algorithms. Up sampling is accomplished by use of an interpolating FIR algorithm; low pass filtering prior to down sampling is accomplished by implementation of a standard FIR algorithm.
6) Two additional C5510 DSK cards were ordered and received this period. Upon receipt the cards were verified for proper functionality by running the developed prototype applications. The new cards came with a newer version of the Code Composer Studio IDE (v3.1). This newer version has been installed on a second PC to provide an alternate development station and to support future segregation of the encoding and decoding phases of the data processing.
7) Finally, time was given to cleanup and document the source files developed to date providing some clarification of what each source file accomplishes and how. Also, documentation of pertinent TI application notes, DSP library routines and various chip support and board support library documents have been organized into a three ring binder for easy reference.
8) The ADPCM library modules have been employed in a 24kHz vocoder application this period.
9) Two approaches have been implemented. The first 3X up samples a 32kHz digitized signal to 96kHz and then 4X down samples it to 24kHz. A low pass interpolative FIR filter is applied to perform the up sampling; pass band and stop band cutoffs and the cutoff ripple parameters are documented in the project. The signal is then companded and encoded according to the G.726 (ADPCM) standard. After decoding, expanding and performing the requisite up/down-sampling the signal is low pass filtered to remove aliasing in the output signal. A second application performs a simple 3x up sampling of a 8kHz digitized signal. As expected, the more robust output is observed in the former approach, but the latter employs less processing load.
10) In other work the encode and decode functions were split into two separate applications for running on individual DSP platforms. Direct transmittal of the (DAC) digital to analog conversion of the encoded signal of one card to the (ADC) analog to digital converter of the input of a connected card does not succeed in maintaining the integrity of the encoded signal. Various attempts to shift the information to upper bits, duplicate the encoded data in adjacent nibbles, or encode the data using (GCR) Group Code Recording encoding methods did not preserve the signal. It has been decided that a transmission scheme incorporating the use of modems will be required to preserve the encoded signal across the transmit function. This will be investigated further next month.
11) A search was conducted to determine the proper procedure to develop a vocoder application to boot and run stand alone on the DSK card when powered up. A literature search on the TI website identified an application notes that discussed use of the ‘C5510 Bootloader and a description of the Hex Conversion Utility from the ‘C55x Assembly Language Tools User’s Guide. Attempts to apply the information obtained from study of these documents failed to

produce an operational stand alone application. A query has been sent to the TI DSP support center.

12) Miscellaneous items accomplished include the installation of a 4-way KVM switch to share one keyboard, mouse and monitor station. Unfortunately it became necessary to upgrade the older PC running Windows 2000 to XP as no previous knowledge of passwords for installed user accounts was available. The second PC also needed to have XP re-installed as efforts to clean up user accounts on that system resulted in boot-up failures. Both systems have Windows XP service pack 2 installed with new user account definitions for all radio lab users.
13) An adaptation of the CVSD voice coding method has been developed for the TI TMS320C5510 DSK platform. This enhances the compliment of vocoders of interest for the radio lab. The current CVSD adaptation operates on voice data sampled at 8 kHz and up sampled to 16 kHz by use of an interpolative low pass FIR filter algorithm. The algorithm allows for specification of the minimum step size and ratio of maximum step size to minimum step size. The number of bits employed in the slope-limiting detection logic is also parameterized to accommodate selection of the common values of N=3 or N=4. Separate encoder and decoder functions have been developed along with definition of CVSD encoder and decoder typedef structures, and initialization routines for acceptance of selected parameters and initializing of internal data values. All functions and typedef structures have been linked into a C5510 library build for easy linking with application software. This vocoder library is referred to as the gd55xCVSD. It currently supports the small memory module builds; both Debug and Release versions of the library have been created. A bootable release version of this application has been successfully programmed into the flash memory of one of the DSK cards for stand alone operation.
14) The DSP support functions that include the re-sampling functions and the audio companding algorithms have been linked into a C5510 library for easy linking with application software development. This library is collectively referred to as the gd55xDSPLib; both Release and Debug versions are provided and support the small memory module build; a large memory model build will be necessary to link the library with existing MELP and ADPCM vocoder applications.
15) It has been determined that a modem link will be necessary to effect transmission of encoded voice data from an encoding-transmit station to a receiving-decoding station. To this end the availability of UART interfaces for ‘C5510 was investigated. Two vendors were identified that provide daughter card UART hardware for the ‘C5510; however, only one was currently supplying daughter cards. An appropriate modem for the application was also researched. One was identified that is reported to work in many OS modes, including DOS, and will function without the need for DTE programming. This modem vendor and the active daughter card supplier information have been recommended to the TM for acquisition.
16) Resources for VoIP encoding and TCP/IP stack software modules compatible with the ‘C5510 DSK platform have been researched as a backup or compliment for the expected SLAB/JVOIPLIB capability from VRSonic. Some vendors are providing technical information and published price lists which will be reviewed and kept on file for future needs as work in VoIP is engaged in the lab.
17) A third PC (Pentium IV, 2.4 GHz) has been configured with Windows XP and the Code Composer Studio v3.1. This PC has been primarily used for the CVSD vocoder development, but can be used as an alternate platform to maintain all of the other vocoder applications and libraries developed to date. The first (and oldest) PC with version 2.2 of CCS will be phased out as a development machine and used primarily for backup and occasional CCS execution station and other organizational tasks.

18) The DSP Global UART daughter cards were checked out using the provided UART driver API and demonstration applications. Once operation was verified the UART interface daughter cards were programmed into two different vocoder projects. One project is a MELP encoder and the other is a MELP decoder. The encoder project MELP encodes a speech signal and transmits it via a direct serial (RS-232) cable link to a second DSK platform hosting the MELP decoder project. The decoder receives the encoded speech signal, decodes it and plays the resultant speech signal over the codec analog line out. To accomplish the synchronization of the decoder with the encoder a two character start sequence ($C0_{16}$, $C0_{16}$) is encoded at the start of every 18 character (left and right) channel data block. A CRC is also computed on the channel data block and appended to the message. The decoder looks for the first start sequence and then reads in the remainder of the message block and discards it; decoding commences on the next message block. The CRC is checked for data integrity. If the CRC is bad, the current channel data is discarded and zeros are fed to the decoder. If the decoder gets out of synch with the encoder, then the decoder resets and looks for the start sequence over again. A modification was made to the DSP Global UART interrupt routine to trigger a software interrupt (SWI) whenever a message block is received. This feature is enabled once the decoder has recognized the first start sequence for reading in subsequent message blocks. A GD version of the library has been created that includes the UART symbology, data structures and API modules. The interrupt service module is presently not part of this library configuration, but is included in the vocoder project for easy referencing of the SWI module for reading of synched message blocks. This library is referred to as the gdUARTlib5510.
19) Programming for the modems was developed this period. Exploration of programming requirements was investigated by first connecting one of the modems to a PC COM port and exercising it using HyperTerminal. After some familiarization working with modems was gained, it became apparent that connecting two modems would require a link that duplicated the load and electrical properties of a telephone line. An online search for direct modem connection methods turned up a circuit description for simulating telephone line characteristics. Once implemented it was then possible to have two PCs connect using HyperTerminal through the modems. HyperTerminal is used to send a command to one modem to dial and the other to answer. This methodology is being ported to the MELP encoder and decoder projects. The encoder will be the answerer (server) and the decoder will be the dialer (client). By means of this, connections have been achieved that permit the MELP encoded data to be transmitted to the decoder and successfully processed there. The algorithm is currently being worked to facilitate successful connections each time and to be able to recover from loss of message block synchronization (restarts).
20) The transmission of MELP encoded speech over simulated phone line was completed this period. The direct cable connection encode and decode applications developed last period were modified to incorporate the modem link. Initialization commands are sent to the respective modems and the decode application performs a dial command while the encode application performs an answer command; either application can be started first. A header block was defined to facilitate synching of the decoder application with the encoded speech stream and a CRC character is appended to each frame of encoded data to insure data integrity across the link. Analog stream is sampled and digitized by the encoder application at 8 kHz, is MELP encoded and transmitted to the modem link at 19.2 Kbaud. The decoder application receives each frame, validates the header and CRC and decodes the data for output to the codec. The direct cable and modem link application pairs both use the same header and CRC data synching and validation algorithms. Two channels are supported. The

encoding applications make use of the on-board switch register for enabling and disabling of channels; disabled channels are processed with a zero digital stream input.

21) Similarly, encode and decode applications were developed to perform transmission of ADPCM encoded speech over a direct serial cable connection. Presently an encode application exists for each of the three sample rates of interest (16, 24 and 32 kHz) and a single decode application has been developed to handle the decoding of all sample rates and speech companding (a-law and µ-law) methods. A mode byte is included in the header of each message block in the data stream that defines the sample rate, companding method and left/right channel processing. When the decoder detects a change in mode, the decoder automatically reinitializes for the new parameters and re-synchs with the encoder-transmitter. Switching of channels does not require reinitializing and data synching. Two encode nibbles are being packed per byte transmitted to reduce bandwidth. Even so, the required baud rate to support 16K ADPCM, two channels, is 230.4 Kbaud; 230400 is the maximum baud rate supported by the UART interface. Consequently, only one channel is supported for transmission of 32K ADPCM; 24K is also only one channel transmission. A modem application will require further reduction in bandwidth. 16K ADPCM will only support further reduction of bandwidth by further packing of the encoded data bits. The encoding applications make use of the on-board switch register for selection of companding method and channel selection.
22) An all-sample rate ADPCM vocoder application was developed to run stand alone. The application will handle 16, 24 and 32K ADPCM encoding and decoding of analog stream input to the on-board codec. The switch register is used to select the desired sample rate and companding method. The processing of either the left or right channel is also selected via a switch on the register. The sampled input is compressed according to the selected method (a-law or µ-law), encoded, decoded and expanded back to 16 bit samples and output to the board codec. This application has been programmed into one of the DSK boards and, together with previous developed MELP and CVSD vocoder applications was used for a project review and demonstration in VOCRES.
23) The source and availability of two more TMS320C5510 DSK boards were obtained and a price quote received from the identified supplier. The quote was passed to the government monitor for purchase. Two additional UART daughter cards were also purchased and received. The UART cards have been configured for ‘C5510 operation, checked out, and placed into service in the lab.
24) The UART interrupt service routine has been returned to the gd55xUART5510 library by including a pointer reference to a (SWI) software interrupt object that the user application defines as the address of the user SWI routine for handling data frame receptions. This serves to make interfacing and development of user UART-dependent applications cleaner and easier.
25) Work has been started to consolidate all the ADPCM encode and transmit applications into one all-sample rate and companding method application similar to the ADPCM vocoder application. This will serve to make ADPCM application support easier and simplify experiment setup in the future.
26) The 16K data rate ADPCM part of the ADPCM universal encoder and decoder applications was selected for adaptation to modem communications. A preliminary analysis of the data rate indicated that it should be possible to support one channel of 16K ADPCM at 38,400 baud; however, the analysis will later prove erroneous. When completed the required baud rate proved to be 57.6 Kbaud. Further analysis, which included asking the support team at Adaptive Digital Technologies, showed that the true data rate was actually 47.5 kbps; packing of the ADPCM codes four per byte was not reducing bandwidth, but only keeping it in check.

Although the applications can communicate over a direct serial link, they can not with a modem link. Broadband server devices that support modem emulation were researched and selected for use in place of the modems. These devices have been placed on order and is hoped that a broadband link can be programmed for ADPCM. It is expected that all ADPCM data rates will be supported with this configuration.

27) The CVSD vocoder application was split into its component encoder and decoder parts with UART interface support for serial link communications. A two channel version that communicates over direct cable link at 57.6 Kbaud was created. Like the original CVSD vocoder application, it samples an analog stream at 8 kHz and up-samples to 16 kHz. The decoder application incorporates a synchronize-to-encoder module that parses the incoming stream for the start header and trailing block which contains a CRC. Once synchronized, a receive threshold count is set that causes a (SWI) software interrupt to be generated signaling the receiving of a complete frame. The SWI handler extracts the encoded data from the message frame for decoding by the DMA SWI handler. A 32 kHz sample rate and subsequent down-sampling to 16 kHz was also investigated; however, the performance of the original sounded superior with better (un-quantified) signal to noise. The direct link applications were adapted with modem control, pared to one channel and successfully tested at 34.8 Kbaud. The CVSD vocoder library was modified this month to include an algorithm check and handing of overflow or underflow conditions if the reference value; the change value is adjusted such that overflow/underflow won't occur. Also a large memory module of the library was generated.
28) Work on the universal ADPCM encoder application for direct cable serial link was completed this period. It supports only one channel for all ADPCM data rates like the universal vocoder application with a like switch register definition for selection of ADPCM parameters. A companion universal decoder was also coded and tested. Both of these applications will easily adapt to the broadband data link configuration to be implemented using the broadband data servers currently on order.
29) Documentation continues to be accomplished in parallel with the development of the vocoder applications. Two sources are being maintained: in source file comments documenting the operation and purpose of each of the application modules and notes written in a project journal.
30) Two Sena PS110 broadband Serial Device Servers were received and implemented this period. Time was given to investigating how to use the PS110 units and then configure them for the task of replacing the traditional lower baud rate modems. The PS110 was selected because of the modem emulation support provided by the unit. Once configured for modem emulation and the required serial communication parameters were defined, the units were deployed in the lab. Although, modem emulation is supported, some changes to the modem commands being used in the traditional modem links was required. The dial command parameters were modified to specify an IP and port number and some other format discrepancies were managed. Once configured and the minor code changes implemented, the broadband link began working as expected. With extended use, however, some transmission issues were manifested. An investigation of the issues was traced to buffer overruns or under runs between the asynchronous nature of the DMA and serial data timing. Double buffering of the serial data helped some, but it was determined that multi-buffering was better suited to managing the data for the two asynchronous events. Logic was added to prevent the buffers from getting out of order. As the multi buffering scheme seems as robust or better than double buffering of ping-pong buffering, and the logic is simpler, this method has been migrated to all of the decoder applications for ADPCM, CVSD and MELP decoding (direct link and modem/broadband links).

31) To investigate the effect (if any) of baud rate on the broadband links, the ADPCM 16 K data rate encoder and decoder applications was made to work at 57600 baud. This was done by packing the 2-bit encoded samples four to a byte and processing one channel. Although not immune to glitches, the slower rate implementation seems to have somewhat better reliability, and has been shown to work well for several hours. It is also believed that the PS110 has a warm-up period after which it performs better. It is recommended that the PS110 units be left on continuously during data collection. Some time was given to investigation of lowering the ADPCM baud rate from 230400 baud to 115200 baud. Reducing the overhead by leaving out up to two characters from the header and trailer would not reduce the bandwidth a sufficient amount so that 115200 baud could support it.
32) During the recent development with the ADPCM applications, a change was designed in the logic that enabled the single channel output to be played binaurally via the onboard codec. For applications that have the capacity to process dual channels, the channel separation is preserved in the playback, but for single channel processing the binaural playback is used. This change has been propagated to all of the applications and new standalone versions of the ADPCM, CVSD and MELP vocoder applications have been flash-programmed in the three DSK boards to be used for data collection.
33) While the Serial Device Servers were still on order, the recent release of SLAB3D incorporating JVOIPLIB VoIP sound sources was installed for testing on the lab PCs. JVOIP was made to function on a local network containing 3 PCs. An unsuccessful attempt was made to get SlabScape, the window form view SLAB3D demonstration program, to allocate and render a VoIP sound source generated from the JVOIPLIB demonstration program. Consultation with the SLAB3D developer did indicate some corrective approaches to try. This work was suspended when the PS110 units were received.
34) Documentation is still being accomplished as the various applications reach completions; source file documentation is being done along with development and usage notes in a project notebook.
35) An effort was made to run encoding and decoding functions, while transmitting the encoded data stream to the encoder over a serial modem link, all operating on one DSK board. This was done to satisfy a request from the branch scientist to compare decoded data with the original data stream. To do so with the existing two board configuration would have required use of RTDX calls and data file management and analysis software. While the single board configuration was achievable, given the dual port UART daughter cards available, it proved problematic in maintaining a constant data stream processing. Real time deadlines with the receiving and decoding of the encoded data stream would occasionally miss. This was observed in the ADPCM encoder decoder application pairs running at 230.4 Kbaud. The application structure was ported to the 16 kHz data rate with quad packing of the encoded samples to allow the much slower 57.6 Kbaud; however, the same receive deadline misses still occurred. The problem was also observed with the MELP encode/decode pair running at 19.2 Kbaud. Here the intensive computation of the encode and decode functions were over taxing the processor capability to meet real-time deadlines. A similar version for CVSD was not pursued.
36) A simplification in the logic for the multi-buffering of the serial data for the modem links was developed this period. In place of the dual ping/pong buffers for receiving of serial encoded data, a multi-buffer scheme which accommodates 3 – 5 buffers (constrained by memory) is used with an arrayed flag corresponding to each buffer. The receive buffer modules fills the buffer and sets the corresponding flag; the data decoding modules processes the data and clears the flag. Each module maintains its own buffer management pointer and the processing module has a “look ahead” feature to look for data if the expected buffer is not currently employed. The logic is simpler than the ping/pong logic and proved more robust in extended running periods.

37) Documentation for all applications has been brought up to date both in the source files and the project notebook. A brief description of each application module and general program flow has been provided along with tables listing the vocoder application for each configuration and identifying key parameters for each (channels supported, data rates, baud rates file names, etc.).
38) A new version (v6.0) of SLAB 3D was received and installed on the lab computers, along with JVOIPLIB which is an open-source VoIP library. Two modules (sending and receiving) are needed to perform the connection; the sender and the receiver can be on the same PC. Two sending applications are available in the release, SLABCall and JVOIPTestUtil; SLABScape (also available) can serve as the receiver application. Using an instance of SLABCall (or SLABCallFile, a variation which can use wave files as a sound source) for every PC in the network, and one instance of SLABScape which is configured to receive each of the available VoIP sources, a full duplex spatialized network can be demonstrated. Each connection must be on a separate IP and port and be a unique session otherwise the various sources will get mixed at the same location. To date this has been accomplished with a three PC LAN.
39) Support was provided as needed for the setup for data collection of performance data utilizing the each of the vocoder applications singly, and then in tandem. A fourth board was programmed with the ADPCM vocoder application so that two data collection stations could be provided for in the tandem mode. One of the vocoders (MELP or CVSD) was found to be generating binaural output by performing individual processing (decoding) of each duplicated data stream producing a slight dither in the output. The application was corrected to duplicate the decoded data stream and output to each channel, thus eliminating the dither.
40) VoIP source allocation support was added to the (IPSS) Internet Protocol Server. The IPSS can act as a VoIP data receiver and render the VoIP data at prescribed locations along with all other supported source types. The IPSS does not send VoIP data and requires that a VoIP caller (send) application be run on the local host or other PC within the LAN. The IPSS user manual has been updated to reflect the management of VoIP data as a sound source. A client program was written to test the VoIP capability in IPSS and currently allocates 3 simultaneous VoIP sources. This program can serve as a prototype for a VoIP communication network demonstration program. In related work, the utility program SlabCallFile was enhanced to allow for browsing and opening of wave files to be used as source data in VoIP connections; the original version only supported hard coding of wave file names. This program supports both wave file and audio device (microphone) data sources and may be applicable for the helicopter brownout study discussed elsewhere in this report.
41) Discussions with the Technical Monitor have identified needs for the experiment to conduct with a VoIP network. Perturbations of the VoIP data stream will probably necessitate the manipulation of the data stream output module in the JVOIPLIB. Work is now planned to investigate the source code released with SLAB 3D to identify areas of interest. Once that has been done discussions with the JVOIPLIB developer may be appropriate.

<u>Intelligibility Measurements of MELP, CVSD, ADPCM and VoIP Vocoding Techniques:</u>

1) Set up Vocoder and ARC-210 radios, in tandem, for speech intelligibility study
2) Monitored study throughout the day and adjusted radio/Vocoder settings to correspond with the different experiment conditions

<u>3D Audio in Helicopter Brownout Effort</u>

1) The feasibility of conducting the helicopter brownout study and demonstration was explored with SIRE facility manager and developer. Discussions focused on the software architecture of SIRE and the data distribution network. Data objects which need to be added to the network to support 3-D Audio have been discussed and identified. A template VSC++ application which runs in the SIRE environment has been received and is being used as a model for the development of a 3-D Audio control node to run in SIRE. The goal of the

control node will be to define locations of the fixed sound sources (ground control, ground outpost, etc.), acquire mobile source locations from the data network, and provide attenuation control. A multi-line text display will be included to provide status and event display.

2) A software application is in development for this study. Work to date has focused on the SLAB interface structure to support the various sound sources; sources for a base station, ground control, wingman and C130 support are being provided. One or more of the sources may be a VoIP while the remaining sources will be ASIO or wave files. Hooks have also been included to update position information for the sources and own ship orientation from the common data area of the SIRE system. A GUI control interface for the audio application has been developed to allow for enabling and disabling of the various sources as well as mute and volume control. This work effort is approximately 40 percent complete.
3) Four new desktop PCs have been received for the Radio Lab. It was determined that the default desktop configuration installed on these systems did not allow for the installation of program development tools like Visual Studio. In order to make the systems useable for program development the systems need to be reformatted and reconfigured for general usage. One of the PCs has been selected for reconfiguration. Visual Studio has been installed along with SLAB 3D (v6.0) and the DirectX SDK. Due to priorities set for the completion of the Dismounted Navigation System, no other work was accomplished for this task. The remaining PCs will be reconfigured during the next report period.
4) The three remaining PCs for the Radio Lab were reconfigured for development use with installation of Windows XP and SP2; other development tools can now be installed as required. One of these PCs (the one with VS 2003 and SLAB) was integrated into the SIRE facility network for use as the audio control workstation in the helicopter brownout study and demonstration. Network communication was verified by using Network Places and Windows Explorer to view and copy files from other network PCs. A missing DLL for the data sharing network (IDATA) was copied to the audio PC in this manner. Running the present version of the audio control software indicated an error condition related to a component of the audio control GUI form and needs further investigation. Further work with this task has been temporarily suspended in favor of work on another task with shorter suspense.

Dismounted 3-D Audio Aided Navigation

1) A software application is being developed to provide 3-D Audio navigation cues for a dismounted (on foot) rescuer in a (CSAR) Combat Search and Recovery environment. The system utilizes a device that report heading information and provides GPS location. The program accepts GPS coordinates of an object or interest and gives audio cues to guide the rescuer to the location. A GUI provides controls to input the target location and displays current information as to location relative bearing to the target and distance. This effort is about 50 percent complete.

2) The software of the dismounted navigation system was completed this period in time for the target demonstration at the Commanders Challenge event. The parsing of the DRM data stream was completed and validated; the Xsens tracker software interface kit was installed and validated. The system was subjected to several evaluations by branch personnel who re-directed some of the effort and helped to make the system more robust and presentable. Clock angle and distance threshold cues were replaced by vocal effort cuing. The head tracking component for the system discarded because it made the system more cumbersome and less intuitive for the user. By working with branch personnel the system was integrated with the Falcon View user environment to support loading of waypoint coordinates directly into the latitude and longitude edit controls on the GUI; however the navigation system was made to input a series of waypoint coordinates from a text file for use in the demonstration. A TDY was conducted to Melrose Bombing Range in NM, near Canon AFB, to support the demonstration of the system during the Commander's Cup challenge event. The system was

successfully demonstrated to about 10 individuals including several combat controllers participating in the challenge.

3) Upon return from the TDY, the system was made to accept target coordinates either from Falcon View inserting them directly into the GUI edit controls, or from a text file that Falcon View inserts coordinates into. Further development of this system is pending new requirements from the branch.

Helicopter Brownout Study

1) Finish the initial development phase of the Audio Control application and prepare for the start on SIRE integration effort.
2) Perform SIRE integration of Audio Control application system.

Wearable Computer Support

PC Based 3-D Audio Rendering

1) An experiment application being developed in the CAVE for SAR scenarios was experiencing performance issues related to sound source rendering using the (IPSS) Internet Protocol SLAB Server. It was determined that the numbers of sources being allocated and rendered were taxing the performance load capability of the host computer. A means of enabling and disabling the sound sources on an "as needed" basis needed to be developed. Since the sources in use were all wave files, it was necessary to give notice to the client application when a source file is complete. A capability to "notify" the client when a wave file is finished playing (single shot only) has been added to IPSS. When allocating a single-shot (non-looping) wave source the client can request notify-when-complete status. The client application may then query the play state of a wave source by sending an "IsDone" command. This capability constitutes an update to IPSS version 2.0.2 and has been verified with a development test client. Test and integration in the CAVE environment has not yet been accomplished.
2) Researched, ordered and received cables for adapting computers to other USB devices
3) Ordered several Wearable Computers from ITRONIX, Inc.
4) Assisted in test of USB port options for the wearable computer systems
5) Researched and ordered several headsets to be tested and utilized with the small wearable computer
6) Ordered more hardware for Wearable Computers from ITRONIX, Inc.
7) Completed and tracked several purchase requests and orders of equipment and materials for this program

Acoustic Signal Control: The NOSH inter-laboratory test in REAT is complete with the following conditions:

1) 1240 3M 1427 Method B (complete)
2) 1241 3M 1427 Method A (complete)
3) 1242 Aearo Peltor H7 Method B (complete)
4) 1243 Aearo Peltor H7 Method A (complete)
5) 1244 Moldex Jazz Band Method B (complete)
6) 1245 Moldex Jazz Band Method A (complete)
7) 1246 Custom Protect Ear dB Blocker Method B (complete)
8) 1247 Custom Protect Ear dB Blocker Method A (complete)
9) 1248 Howard Leight Air Soft Method B (complete)
10) 1249 Howard Leight Air Soft Method A (complete)
11) 1250 Aearo Peltor EAR Classic Method B (complete)
12) 1251 Aearo Peltor EAR Classic Method A (complete)

13) 1275 Gentex HGU-55P Mask, Visor, Oregon Aero Zetaliner, David Clark ANR earcups (Active) + Undercut Comfort Gel Earseals + Westone Labs ACCES Gen4 Aircrew (complete)
14) 1276 Gentex HGU-55P Mask, Visor, Oregon Aero Zetaliner, David Clark ANR earcups(Passive) + Undercut Comfort Gel Earseals + Westone Labs ACCES Gen4 Aircrew (complete)
15) 1282 Gentex HGU-56P Visor, Face shield, air bladder, Active Xtreme passive earcups w/gel seals (complete)
16) 1283 Gentex HGU-55P, mask, visor, Oregon Aero Zetaliner, Active Xtreme 56 ANR (Active) earcups, w/oval shims, gel earseals + Westone Labs ACCES Gen4 Aircrew (complete)
17) 1273 Gentex HGU-55P Mask, Visor, Oregon Aero Zetaliner, Gentex ANR earcups (Active) + Westone Labs ACCES Gen4 Aircrew (cancelled)
18) 1280 Gentex HGU-56P Clear Visor, TPL liner, David Clark ANR earcups (Active) + Undercut Comfort Gel Earseals + Westone Labs ACCES Gen4 Aircrew (complete)
19) 1284 Gentex HGU-55P mask, visor, standard earcups + OA Zetaliner & Softseals + Aearo EAR Classic 50 percent insertion, (complete)
20) 1287 Gentex HGU-56P clear visor, TPL liner, David Clark ANR earcups (passive) & undercut earseals + Westone Labs ACCES Gen4 Aircrew (complete)
21) The updated HPD Attenuation List was posted to the government website.
22) The new room microphone in REAT is being used for daily calibration on the existing system.
23) The room microphone in REAT was damaged, which prevents performing daily calibration. Testing continues however, and a different type of replacement microphone is on order. This new mic is needed for the full REAT system upgrade still in progress.

MIRE

1) Sound Pro microphones replaced the Knowles 1785 in the MIRE facility. A 10 subject test was conducted on the previously tested David Clark H1076-XL to validate these new mics.
2) 1224 Gentex HGU-55P Mask, Visor, Oregon Aero Zetaliner, David Clark ANR earcups+ Undercut Comfort Gel Earseals (complete)
3) 1234 Gentex HGU-56P Visor, Face shield, air bladder, Active Xtreme 56 ANR earcups, gel earseals (complete)
4) 1235 Gentex HGU-55P, mask, visor, Oregon Aero Zetaliner, Active Xtreme Stealth ANR earcups, leather seals, triangle shims (complete)
5) 1236 Gentex HGU-55P, mask, visor, Oregon Aero Zetaliner, Active Xtreme 56 ANR earcups, gel earseals, oval shims (complete)
6) 1233 Gentex HGU-56P Clear Visor, TPL liner, David Clark ANR earcups+ Undercut Comfort Gel Earseals (complete)
7) Two locked storage cabinets were added to room 1-24 for storage of HPDs. This room was cleaned out, vacuumed, and light bulbs were replaced.
8) The REAT, MIRE, audiometric chamber, and control areas were vacuumed and dusted.

The Earplug Material and Construction test is underway in REAT with the following conditions:

1) 1255 Westone Labs Solid Soft Silicone (8/10)
2) 1256 Westone Labs Solid Hard Silicone (7/10)
3) 1257 Westone Labs Solid PVC (complete)
4) 1258 Westone Labs Earphone Hard Silicone (7/10)
5) 1259 Westone Labs Earphone Soft Silicone (8/10)
6) 1260 Westone Labs Earphone PVC (complete)
7) 1261 Westone Labs Earphone & Mic Hard Silicone (7/10)
8) 1262 Westone Labs Earphone & Mic Soft Silicone (8/10)
9) 1263 Westone Labs Earphone & Mic PVC (9/10)

The Acousticom test is underway in both REAT and MIRE with the following conditions:

1) 1264 (REAT) Gentex 55P-lw w/mask and visor, noHose, OA Zetaliner+Acousticom H154ENC (9/10)

2) 1222 (MIRE) Gentex 55p-lw w/mask and visor, noHose, OA Zetaliner+Acousticom H154ENC (8/10)
3) The Gen4 test (1252) in REAT on hold until we can get more helmet sizes and earplugs
4) 1252 Gentex 55P-lw w/mask, visor, and Hose, OA Zetaliner and Softcups+ACCES Gen4 (10/20)
5) Audiograms are not being conducted in the new chamber. Calibration of GSI and Earscan audiometers was certified by Gordon and Stowe

BAM Lab/BATMAN

BAO Batman

REAT

1) 1285 TI miniTAC complete
2) Modified and enhanced photos of BAO equipment

Navy HPD Measurements

REAT

1) In anticipation of the Unit Compliance Inspection, the former electronics lab cabinets, as well as the equipment room racks and wood shop, continue to be consolidated and organized. Content sheets and labels were applied to cabinets and shelves.
2) 1272 Gentex HGU-56P Clear Visor, TPL Liner + New Dynamics Sound Guard Two Color(complete)
3) 1274 CEP, Inc. “Mini-CEP” CEP505-C11 (19/20)
4) 1270 Gentex HGU-56P Clear Visor, TPL Liner + Westone Labs ACCES Gen4 Aircrew (complete)
5) 1274 CEP, Inc. “Mini-CEP” CEP505-C11 (might be restarted due to part variation)
6) 1279 Silynx (complete)
7) 1281 Creare STTR without face shield + Westone Labs Solid Vinyl PVC plug (complete)
8) 1286 JSF ANR ACCES unpopulated (complete)
9) 1288 Aegisound Max25/40 + JSF ANR ACCES unpopulated (9/10)
10) 1290 JSF ANR ACCES unpopulated (version 2) (3/10)
11) 1291 Aegisound Max25/40 + JSF ANR ACCES unpopulated (version 2) (3/10)
12) 1292 Silynx Quietops w/Comply Canal Tips (10/20)
13) Experiments 1288, 1290, and 1291 will be continued on a CRADA.
14) Eight subject panel members and twelve ad hoc subjects were scheduled to participate in nine studies during this reporting period. Earmolds were made for different sets of custom earplugs.
15) Ten subject panel members and nine ad hoc subjects were scheduled to participate in ten studies during this reporting period.
16) One ad hoc subject was scheduled and had earmolds made for custom earplugs.
17) Ten subject panel members and seven ad hoc subjects were scheduled to participate in eight studies during this reporting period.
18) Two ad hoc subjects were scheduled and received hearing tests.
19) Two ad hoc subjects were scheduled for earmolds for custom earplugs.
20) Three studies were completed during this reporting period: 1315 CEP Inc. CEP505-C11-V (vented) with Hearing Components Comply Canal Tips; 1317 Westone Labs ACCES Gen5 Groundcrew; 1318 Westone Labs ACCES Gen5 Aircrew; (complete) (complete Data collection is on schedule.
21) Five studies were completed during this reporting period: 1315 CEP Inc. CEP505-C11-V (vented) with Hearing Components Comply Canal Tips; 1316 SureFire EarPro Sonic Defenders EP3 Unstoppered; 1323 Racal Acoustics Raptor (passive); 1324 MSA Sordin (passive); 1325 Peltor Comtac w/pp403 gel earseals.

22) Thirteen acoustic subject panel members and seven ad hoc subjects were scheduled to participate in 14 studies during this reporting period.
23) Three sets of Gen 5 Access earplugs were received and subjects were scheduled.
24) Eight attenuation studies were completed during this reporting period:
    a) 1334 Aero Ear Combat Arms Double Ended (green end inserted in left ear)
    b) Next Link Invisio Pro Digital Standard (right ear);
    c) 1335 Red Tail Hawk Custom Earshell;
    d) 1336 Red Tail Hawk Headset/Custom Earshell;
    e) 1337 Red Tail Hawk Headset;
    f) 1338 Sennheiser SLC110L (port open) with triple flange plug;
    g) 1339 Various HGU-25 with goggles with Safety Direct Aural Protector plus CEP505-C11 (mini CEP no vent) with Hearing Components Comply Canal Tips;
    h) 1340 MSA Sordin Neck (passive test); 1341 Sennheiser WACH 900 (passive test)

MIRE

1) 1226 VSI HMD ANR helmet (complete)
2) 1227 VSI HMD ANR helmet + Randolph Engineering HGU-4/P Sunglasses (complete)
3) 1231 VSI HMD + BT2000 Series 140 sunglasses, (5 subjects, complete?)
4) 1232 Gentex HGU68P w/visor, TPL Liner, VSI ANR, (complete)
5) 1237 Safety Direct tension rig, David Clark shaped cups, straight view, 14 Newtons (complete)
6) 1238 Safety Direct tension rig, David Clark shaped cups, 45° right, 14 Newtons (complete)
7) 1239 Safety Direct tension rig, David Clark shaped cups, straight view, 25 Newtons (complete)
8) 1240 Safety Direct tension rig, David Clark shaped cups, 45° right, 25 Newtons (complete)
9) 1241 Safety Direct tension rig, David Clark standard cups, straight view, 14 Newtons (complete)
10) 1242 Safety Direct tension rig, David Clark standard cups, 45° right, 14 Newtons (complete)
11) 1243 Safety Direct tension rig, David Clark standard cups, straight view, 25 Newtons (complete)
12) 1244 Safety Direct tension rig, David Clark standard cups, 45° right, 25 Newtons (complete)
13) Six subject panel members and two ad hoc subjects were scheduled to participate in three studies during this reporting period.
14) Fourteen Summer Panel members participated in the SOCOM Speech Intelligibility Study in MIRE, in which speech intelligibility of twelve devices was measured.

BAO Speech Requirements

1) Testing of the TAC plugs with the speech recognition system has slipped pending a resolution of the TAC box hardware issues.
2) Integrated speech demonstrations will be given to operators at Winter Camp 2006. Preparations are on schedule.
3) Work on the speech interface for the UAV Targeting Tool continued to hold. SRA International will update their code using a speech API sometime later this year. Current demonstrations of the speech technology will continue to use the “AutoIt” solution.
4) All speech recognition control for the Batman program was reviewed in preparation for demonstration at Winter Camp 2006. The demonstration laptop was updated with new versions of Bareback and the UAV Targeting Tool. The BAOTalk program, AutoIt functionality, and all associated grammar files and documentation were updated to work with the new program versions.

5) Dry runs of the speech recognition functionality and other Batman capabilities were accomplished in the Bamlab prior to traveling to Winter Camp. Speech recognition demonstrations at Winter Camp were very successful.
6) No feedback from AFRL/MN has yet been received on their use of the speech interface with the micro air vehicle.
7) Work on the speech recognition API to support the upcoming SOTACS III contract is again on hold until SRA is ready to start speech integration for their products. Work on the Army FFW version of the speech-enabled BAO tools progressed in preparation for demonstrations to various FFW and other program office personnel.
8) Flight test activities using the VISTA version of the Dynaspeak recognizer were supported. Flight test data and recordings were received from the Test Pilot School at Edwards AFB. Data analysis has begun.
9) The WCAS project code was developed, integrated, and tested.
10) Work continued on the development of the System Voice Control (SVC) code in support of the Army's Future Force Warrior (FFW) and Air Force BAO programs. Weekly Technical and status telecon meetings were held with the entire FFW technical team. Preliminary draft Use Cases developed by SRA were reviewed, and all proposed SVC functionality was collated into a preliminary interface design document. Support for dynamic grammars was added to the FFW grammar and tested with a version of the Waypoint Editor, which was modified to support this addition.
11) Preparations were made to support the March 5-9 Software Technical Exchange and Software Integration meeting to be held at Fort Monmouth, New Jersey. A PowerPoint presentation and briefing was developed which outlines the soldier training suggested for the SVC program. The briefing will be presented and reviewed at the meeting.
12) Weekly Technical and status telecon meetings continue held with the entire FFW technical team. Software work continued on the development of the System Voice Control (SVC) code in support of the Army's Future Force Warrior (FFW) and Air Force BAO programs. The SVC code structure was redesigned to make meet the requirements of both the FFW and BAO programs. The SVC interface design specification was delivered to SRA, along with the SVC package, including the Waypoint Editor, and fully functioning FFW grammar.
13) A trip was taken to Fort Monmouth, New Jersey to integrate the SVC code with other FFW Software components. A briefing was given to outline the soldier training suggested for using the SVC program.
14) Software work continued on the development of the System Voice Control (SVC) code in support of the Army's Future Force Warrior (FFW) and Air Force BAO programs.
15) Relatively minor SVC code changes were made to support the upcoming Army's FFW On-The-Move (OTM) exercise at Fort Dix.
16) Meetings were held with SRA on the interface design specification, Waypoint Editor and FFW grammar.
17) A trip was taken to Natick, Mass to support an Army FFW meeting. This meeting served as a "hotwash" for the 2007 efforts. In addition, plans for 2008 support were discussed.
18) BAO program technical status meetings were held. Development of the BAOTalk Spiral 2 release continued. Additional functionality was added to the speech interface to FalconView with the incorporation of a "panning" command. Operators can now move the map center location by saying "pan left 2 degrees", "pan south 1 degree", etc. Work on the Mini Air Vehicle (MAV) speech interface also progressed. An interface control specification was worked out with AFRL/MN and Applied Research Associates, Inc. This specification outlines a formal communication protocol for data transferred between the MAV OCU and BAOTalk.
19) Traveled to Fort Walton Beach, FL to attend the BAO Summer Camp at Eglin Field 6. Demonstrations of the BAO speech recognition system were given to various participants.

Working technical discussions were held with various other BAO team members during the Summer Camp to address speech interface issues.

20) Work on the speech recognition API to support the upcoming SOTACS III contract was initiated. The speech recognition BAO laptop used for demonstrations was updated with current BAO software versions. Additional speech commands were added to support demonstrations to various customers. The “Barebones” version of Bareback was configured with new grammars and functionality to support Army requirements.
21) Audio analysis of Dynaspeak recordings was started.
22) A TDY was taken to support the paper selection process for the October NATO symposium on human factors issues in autonomous military vehicles. Speech recognition technology is expected to have a major role in the solution of future operator interface challenges for these systems.

Target Acquisition Support, Wearable Computing Support, and BAO BATMAN Program Coordination: BAO BATMAN Program Coordination is a 6.3 program to enhance Warfighter capabilities by developing technically advanced tools that are human centered. Target Acquisition Support is focused on providing the Warfighter with human=centered display system that assist in target recognition. The Wearable Computing Support project is a Congressional Add that is developing miniaturized table computers for the operational community.

1) Task 1: Target Acquisition Support
    a) Attended summer camp 2005 at Ft. Walton Beach, FL
    b) Worked on the feasibility of using a radio pod to transmit data wireless on the battlefield
2) Task 2: BAO BATMAN Program Coordinator
    a) Coordinated and attended various status meetings for subordinate efforts in this program
    b) Tracked progress of subcontracting efforts
    c) Tracked and updated program scheduled and milestones
    d) Tracked and updated program spending burn rates
    e) Attended summer Camp 2005 and demonstrated communications equipment
3) Task 3: Wearable Computing Support
    a) Continued discussions with the para-rescue community and discussed the feasibility of adding 3D audio technology to the CSAR mission

Informational and Energetic Making

1) Repaired audio communications system in X-facility better facilitate lab audio connection and use.
2) Cleaned up old projects out of the VOCRES facility and ordered new equipment
3) Rewired and repositioned the ALF Head tracker into another room for better noise abatement in the ALF chamber
4) Revamped and equipped the ALF Chamber Demonstration with new Cordless RF Headsets to allow better show and performance of the ALF Chamber Demo
5) Revamped and equipped the ALF Chamber Demonstration with new Cordless RF Headsets to allow better show and performance of the ALF Chamber Demo
6) Rewired several sets of headsets for testing in this program
7) Ordered more clay / putty for all the speakers in the ALF Chamber
8) Completed and tracked several purchase requests and orders of equipment and materials
9) Modified microphone setup for calibration purposes
10) Manufactured microphone setups and power boxes for ALF tests under this program

NetCentric Exp

1) deviceLocalization is an audio localization experiment testing subjects’ ability to localize ¼ second and continuous sound clips while wearing a variety of hearing devices.
2) Ear plugs and ear plug systems being tested are Combat Arms Ear Plugs (CAEP), Mini DTAC custom plugs, Nacre foam plugs, Peltor Comtac headset, Peltor Sportac headset, SensiMetric custom plugs, Silynx foam plugs, Bose headset, CEPS foam plugs, MSA Srdn headset, and SoComX custom plugs.

3) 9 subjects were tasked to run 26 blocks (2 blocks for each device plus 2 blocks with an open ear condition). Each block contains 180 trials.
4) Study completed

Enhanced MMC Monitor Software Development

1) Received a copy of the WCAS software, DIS log player software and DIS log files, and simple MMC monitor software from RHCP and worked with RHCP to get it installed and working on a target laptop PC. In the performance of this work it was determined that the laptop PC OS needed to be reconfigured without desktop management to allow WCAS to run and to allow for program development software installation. The re-installation of the laptop OS and installation of the VS development suite was accomplished. WCAS operation was then verified on the laptop PC.
2) An enhanced version of the MMC Monitor software developed by RHCP is in development to allow operator interaction and control of real time and captured audio. Controls have been added to the original form to provide spatial presentation control of real time speech with mute and volume control capability, playback of captured speech, and filtering of content based on frequency or other Entity ID. The enhanced monitor, or MMC Monitor Plus, software is about 50 percent complete. The GUI design and layout, including the selection and playback of captured speech, is all but complete. The work of commanding the IPSS to allocate and render DIS channels is in the early stage of development as the version of SLAB (6.0.1) which supports DIS sound sources has just become available.
3) An intermediate (IPSS) Internet Protocol SLAB Server Manager has been developed to manage multiple client connections to the IPSS, which renders the real time speech channels. Each MMC Monitor is being designed to present a particular DIS channel and use SLAB, via the IPSS, to spatially present that channel at a selected location. Since the IPSS only accepts connection to one client at a time, the IPSS Manager will handle the various MMC Monitor Clients, passing requests from the monitor clients to the IPSS and routing the responses back to the appropriate monitor client. The IPSS Manager also handles commanding IPSS to set up the SLAB environment, including the HRTF dataset load, and render sample rate specification. The IPSS has been upgraded to accept commands to allocate DIS sound sources and is linked with a pre-release version of SLAB (v6.0.1) libraries. The development effort of the IPSS Manager server is about 80 percent complete with some final tweaking of the message handling and routing yet to be accomplished.
4) A speech activity indicator was incorporated into the MMC Monitor Plus GUI this period. The real-time speech position indicator bullet is made to display red whenever activity is detected on the selected DIS channel. It was necessary to modify the (IPSS) Internet Protocol SLAB Server program to implement VU monitoring functions on the DIS source when it is allocated. Every 250 ms, or so, the IPSS will sense the state of the DIS channel “VU meter” to see if there has been a change in state. When it goes from off to on, or on to off, it sends an unsolicited message to the client manager. The client manager routes the message to the appropriate MMC Monitor client based on the DIS source number. A separate port and supporting software modules, for handling of unsolicited messages was added to the MMC Monitor clients and the IPSS Manager client/server program. The IPSS now expects two connections to be made per client; one for normal communication and one for unsolicited messaging. This functionality needs to be added to the IPSS documentation.
5) In addition to the voice activity indicator, the MMC Monitor GUI header background color will change to a dark green color to indicate that there is a pending speech transcription. When the transcribed speech is presented, the header color reverts to the standard Windows control gray. However, since the system has no way of discerning if the current voice activity state change is part of the same utterance, or a new one, some pending transcription states go uncolored.
6) The monitor program has been enhanced to recognize ‘bullseye’ and ‘BRAA’ keywords in the transcription to attempt rendering of the subsequent words in military brevity format; call sign and several other acronyms are also being recognized and made to display with uppercase letters. A simple text search has been implemented to allow the user to search back for the last occurrence of a word or phrase. The containing item, if found, is highlighted and auto-

scrolling is suspended. Subsequent searches of the same keyword or phrase will look for the next last occurrence, and so on; new searches always start from the last transcribed item.

7) The text is now being displayed with the font Courier New which provide a more uniform and predictable text length. The item height of each text line is now easier to determine and program. This has resulted in utterance transcription always being displayed properly formatted and complete as provided by the transcriber function.
8) An update of the WCAS startup procedure has been received from the speech analysis group. The new procedure is more robust, having corrected the load DLL errors and affords more complete transcription of long (> 30 seconds) utterances. The startup of the MMC Monitor Plus program has been integrated with the new procedure and has been used for initiating up to 6 instances of the monitor software. An error that was causing some transcribed speech wave files to be placed improperly in the file structure was reported to RHCP and fixed.
9) The IPSS Manager client/server program has been observed to be much improved since the client tracking methodology added and reported last month was implemented; however, startup of many (6 or more) instances of the MMC Monitor Plus clients can, at times, generate loss of client and message association. A simple delay of the MMC Monitor client request for source presentation and positioning commands has show to arrest the problem. This is easily implemented by basing the delay amount of each client request on the source allocation number returned to it. Progressive one second delays in each of the MMC Monitor clients provide successful startups.
10) Three in-house demonstrations of the MMC Monitor system for individual representative of the target user community have been conducted this month and participation in the debrief period following the demonstrations was included.
11) A DIS transmit capability has been added to the MMC Monitor Plus GUI application tool. Although more work is needed to fully manage the capability, the transmission can be heard at the users, or other workstation, when the appropriate DIS channel is monitored; transcription of the transmitted speech can be seen in the speech to text display. Also added the period was a guard against allowing the user to select a DIS channel that is already mapped by another instance of the application running at the same workstation which would otherwise SLAB rendering to fail. Tool tips have also been added to inform naive users of the various functions of the GUI controls; notification of when auto-scrolling becomes disabled is also provided as a tool tip and the play control buttons are highlighted in red.
12) A simple search capability has also been added that permits searching of the transcribed text for any keyword desired (case sensitive) from the last transcribed entry to the first; subsequent searches of identical text will find the next subsequent occurrence. New searches always start from the last transcribed item. When searches reach the top a message box is displayed and then the search will ‘wrap’ upon subsequent clicking of the search command button.
13) Other enhancements include the population of the information boxes with frequency and call sign data (if known) associated with the selected DIS channel, the Talk button is now a true PTT (click and hold) function, and the text transcription pending state display now functions independent of the text filter setting.

14) Two other laptop computers were configured and prepared to run the MMC Monitor Plus software tools as satellite stations (work stations that do not host the WCAS speech transcription software). The laptop PCs had their OS software re-installed to remove the desktop configuration software which comes pre-installed. This reconfiguration allows the Web server application to run without hindrance from the firewall software activated by the desktop configuration monitor. Once configured, the SLAB software server (IPSS) and the IPSS Manager client/server software and the MMC Monitor software was installed and configured. A router was obtained and used to network the three laptops and assign IP addresses by DHCP. Tests were conducted to validate the operation of the networked computers. Trial runs showed that the MMC Monitor software needed to be rebuilt to accommodate the IP addresses assigned by DHCP, instead of local host, so that the chat monitor program instances could find the playback wave files located on the hosting laptop.

The testing also indicated the need to assign a unique DIS channel ID for PPT transmitting at each laptop station. Pre-conference in-house demos were conducted to help validate the proper operation of the networked system. A TDY was conducted to the AWACS / AEW&C Program Management Review (A2PMR) conference in Seattle, WA where the (MMC) Multi-Modal Chat Monitoring tool was demonstrated as a side session event on two successive days. The system performed well at the conference. The attendees were given informal one-on-one briefing about the system operation and goals and afforded hands-on demonstration time. Participants were encouraged to ask questions and offer feedback; pro-active attendees who participated in the demo were asked to complete surveys.

CSAR Demonstration Software for JOSHI

1) Provided consultation for SLAB integration into the CSAR software being developed for the JOSHI visual laboratory at Wright State University

Acoustic Signal Control

1) Ordered modification kits for P-56 helmet ACCES testing
2) Built an ear plug power and amplification box for upcoming testing on helicopters
3) Discusses new engineering and procurement techniques to bring about the new Active Noise Reduction ANR Custom fit communications plug system to the USAF aircraft cockpit
4) Completed ACCES to David Clark and Bose communication headset modification procedures for Air Combat Command Life Support personnel

RHCB Permanent Subject Pool Management: Subject Panel Management for studies occurring within RHCB labs.

1292 Silynx with Comply Tips

1) Four subject panel members participated in a study in the REAT facility, in which the attenuation of Silynx plugs with Comply tips was measured.
2) Data collection has been completed for ten subjects.
3) Four subject panel members and six ad hoc subjects were scheduled to participate in a study in the REAT facility, in which the attenuation of Silynx plugs with Comply tips was measured.

1295 Combat Arms French Single Ended Earplugs

1) Nine subject panel members and one ad hoc subject were scheduled to participate in a study in the REAT facility, in which the attenuation of Combat Arms French Single Ended earplugs was measured.
2) Data collection has been completed for this study.

1296 JSF ANR 56P Helmet and Mini CEPs (non-vented)

1) Six subject panel members participated in a study in the REAT facility, in which the attenuation of JSF ANR Acces (unpopulated) with Aegisound Max 25/40 Earmuffs was measured.
2) Data collection has been completed.

1297 Peltor Comtac Muff

1) One subject panel member was scheduled to participate in a study in the REAT facility, in which the attenuation of a Peltor Comtac Muff was measured.
2) Data collection has been completed.

1298 JSF Medium Gray Active Extreme Custom Plug

1) Six subject panel members were scheduled to participate in a study in the REAT facility, in which the attenuation of JSF medium gray Active Extreme Custom plugs were measured.
2) Data collection has been completed.

1299 JSF Hard Brown Active Extreme Custom EarPlugs

1) Six subject panel members participated in a study in the REAT facility, in which the attenuation of JSF medium gray Active Extreme custom plugs was measured.
2) Data collection has been completed.

3D Audio Chamber Studies: Subject panel availability and overall operation was monitored for the following studies:

1) **FiltCRMVal**:
    a) Assessed intelligibility of speech that was filtered into multiple bands with masker in overlapping and non-overlapping bands.
    b) Evaluated target identification when target and masking talkers were selected after filtering, so as to minimize spectral overlap.
2) **TMaxMMin**:
    a) Evaluated listeners' ability to track a target and / or cancel out a masker when the apparent location changed with some probability.
    b) Assessed target identification when the target talker was split and presented at two different spatial locations and compared it to conditions when the masking talker was spatially split.
3) **Dichodetectexp2**:
    a) Measured listeners' detection threshold to judge if the target talker was male or female with and without a contralateral masker.
    b) Audio threshold experiment testing subjects' ability to determine the gender of a talker while other stimuli are presented.
4) **Dichodetectexp3**:
    a) Measured listeners' detection threshold to judge if target talker was forward or reverse speech with and without a contralateral masker.
    b) Audio threshold experiment testing subjects' ability to determine orientation of a talker's voice (forward or backward) while other stimuli are presented.
5) **Whisper2:** Evaluates target intelligibility with multiple whispering talkers compared to normal talkers, in order to assess target segregation efficacy in situations where talkers are required to be unobtrusive.
6) **Grouping_control_3talker:** Assesses if the presence of a call sign aided target identification with artificial speech signals, where segregation was found to be difficult.
7) **SpeedCP2:** Assesses the influence of rate of speech on target segregation in a multitalker listening task (both speech and noise maskers) at low signal-to-noise ratio.
8) **Sparse_modtype:** Examined target identification when target figure-ground contrast was either sparse or densely distributed in spectro-temporal bins as a function of signal-to-noise ratio.
9) **Sparse_ITD2:** The study measured the effectiveness with which interaural phase delay segregates a target signal from the background.
10) **CRM Studies** which measure the intelligibility for two types of synthetic CRM phrases in the presence of noise or other interferers.
11) **Scaling Studies** which evaluate the influence of *a priori* knowledge about the characteristics or content of the maskers or the target speech signal on a listener's ability to extract information from the target speech signal.
12) **Spatial Adaptive Studies:** which assess the masking release obtained by separating the target and masker in space.
13) **Spatial Verify Studies:** which assess the effectiveness of head related transfer functions (HRTF) in spatial separation of target and masker.
14) **Alflocalization 2**
    a) Open ear audio localization experiment testing subjects' ability to localize ¼ second sound clips and 30 second continuous sound clips.
    b) 8 subjects were tasked with 9 blocks of ¼ second sound clips and 3 blocks of 30 second continuous audio sound clips. (* 1 block = 79 trials)
15) **Bandslocalization b**
    a) Open ear audio localization experiment testing subjects' ability to localize ¼ second sound clips at varying bandwidths.

b) 10 subjects were tasked with 9 blocks of ¼ second sound clips. (* 1 block = 79 trials)

16) **Bandslocalization 3b**
    a) Open ear audio localization experiment testing subjects' ability to localize ¼ second sound clips at varying bandwidths. Each subject also had to center their head to a predetermined center speaker (#273) between each trial.
    b) 10 subjects were tasked with 9 blocks of ¼ second sound clips. (* 1 block = 79 trials)
17) **Grouping Studies** which address questions about the relative salience of several cues such as onset, fundamental frequency, common modulations and spatial location, and target segregation in multi-talker listening tasks.
18) **MRT AngleTesting Studies** which evaluate the extent of visual contribution (speech reading) in a speech intelligibility task as a function of viewing angle.
19) **Scaling Studies** which evaluate the influence of a priori knowledge about the characteristics or content of the maskers or the target signal on a listener's ability to extract information from the target speech signal.
20) **Gun Exp studies** which evaluate the effectiveness of a transparent hearing protection device by requiring the subjects to localize and identify a target phrase in the presence of gun fire.
21) **Cueing studies** which evaluate the ability of listeners to detect and localize a target phrase which could be one of the following: forward PB words, Reverse PB words, forward environmental sounds and reverse environmental sounds. The effectiveness of cueing will also be assessed by the presentation of a precue or a postcue.
22) **Tanya studies** which asses the identification performance of listeners in the presence of two maskers which are 1) normal speech maskers, 2) Fo maskers and 3) Sineband maskers.
23) **Third Talker studies** which evaluate the effect of a similar vrs a non similar masker on target intelligibility.
24) **CRM_Detect studies** which evaluate if detection thresholds differ as a function of the tasks that the listeners were required to do (for example, detect the presence of a target vs. detect if the target is forward or reversed).
25) **Control_Dicho Detect studies** which assess detection thresholds for a wide variety of tasks tested in CRM detect with and without a contralateral masker, and as the nature of the contralateral masker varies.
26) **Eavesdrop studies** which explore the listeners' ability to detect call-back errors with two dyads (4 talkers) in a spatialized vs. non-spatialized listening condition.
27) **Bands studies** which evaluate the psychometric functions for two kinds of target signals: normal speech and filtered speech, in the presence of two other similar maskers.
28) **FiltCRMVal**: Assessed intelligibility of speech that was filtered into multiple bands with masker in overlapping and non-overlapping bands.
29) **TMaxMMin**:
    a) The study evaluated listeners' ability to track a target and / or cancel out a masker when the apparent location changed with some probability.
    b) Assessed target identification when the target talker was split and presented at two different spatial locations and compared it to conditions when the masking talker was spatially split.
30) **Dichodetectexp2**: Measured listeners' detection threshold to judge if the target talker was male or female with and without a contralateral masker.
31) **Dichodetectexp3**: Measured listeners' detection threshold to judge if target talker was forward or reverse speech with and without a contralateral masker.
32) **DichodetectID4:** The study evaluated the listener's detection thresholds for identifying a target signal in the presence of ipsilateral and contraleteral interferers, both noise and speech. DichodetectID4 study was completed.
33) **Dichodetect_2afc:** This study tested the subjects' ability to detect a variety of tones, sounds or words while another stimulus was presented.
34) **Whisper2**: Evaluates target intelligibility with multiple whispering talkers compared to normal talkers, in order to assess target segregation efficacy in situations where talkers are required to be unobtrusive.

35) **Grouping_control_3talker** which assesses if the presence of a call sign aided target identification with artificial speech signals, where segregation was found to be difficult.
36) **SpeedCP2** which assesses the influence of rate of speech on target segregation in a multitalker listening task (both speech and noise maskers) at low signal-to-noise ratio.
37) **Sparse_ITD2:** The study measured the effectiveness with which interaural phase delay segregates a target signal from the background.
38) **Sparse_ILD2:** The experiment measured the ability of the listeners to segregate a target signal on the basis of a binaural interaural level difference cue. **SpeedCP:** Evaluates listeners' performance in multi-tasker listening tasks with all signals (target and masker) being time-compressed or time-expanded.
39) **PsychoIM Series:** The goal of these experiments was to obtain detection thresholds for a brief sinesoidal tone in the presence of complex multi-tone maskers as a function of signal-to-noise ratio. Three types of maskers were generated and thresholds will be obtained in three different experimental conditions that varied in degree of uncertainty of the target level and / or the masker level.
40) **Sparse Series:** The Sparse series of experiments evaluates the effectiveness of various monaural and binaural cues in the segregation process. In a task that is analogous to visual figure-ground judgments, target signals were generated which differed from the background by only one cue, such as level, interaural time difference, interaural level difference etc.
41) Ran subjects and participated as a subject in the experiments listed below.
    a) Dicodetect_2afc
        1. dicodetect_2afc is an audio threshold experiment testing subjects' ability to detect a variety of tones, sounds or words while another stimulus is presented.
        2. Subjects ran the experiment in an isolation booth while wearing Gen4 Access custom ear plugs.
        3. 8 subjects were tasked to run 10 blocks with a varying number of trials.
        4. Study began on February 26, 2007
        5. Study completed on March 6, 2007

    b) EqualizationFrameSW2
        1. Audio experiment testing the fix and refit of a headset versus custom ear plugs. Subjects are tasked to equalize a sound source between the left and right ear.
        2. Subjects completed the task wearing either a set of SensiMetric custom plugs or a Sennheiser HD 520 II headset.
        3. Study began on February 5, 2007
        4. Study completed on March 28, 2007

    c) DicodetectID4_misc
        1. dicodetectID4_misc is an audio threshold experiment testing subjects' ability to identify a talker spoken color and number combination in the presence of noise and/or other talkers.
        2. Subjects ran the experiment in an isolation booth while wearing Gen4 Access custom ear plugs.
        3. 8 subjects were tasked to run 5 blocks with a varying number of trials.
        4. Study began on March 5, 2007
        5. Study completed on March 16, 2007

Subjects completed Dicodetect_2afc, dicodetectID4_misc, and equalizationFrameSW2.

MIRE Facility:

1) Fourteen Summer Panel members participated in the SOCOM Speech Intelligibility Study in MIRE, in which speech intelligibility of twelve devices was measured. Data collection was completed.

ALF (Audio Localization Facility)

1) Dicodetect_2afc

a) dicodetect_2afc is an audio threshold experiment testing subjects' ability to detect a variety of tones, sounds or words while another stimulus is presented.
b) Subjects ran the experiment in an isolation booth while wearing Gen4 Access custom ear plugs.
c) 8 subjects were tasked to run 40 blocks with a varying number of trials.
d) Study began on October 5.
e) Study completed on November 9.
f) Subjects have completed the dicodetect_2afc experiment.

2) Dicodetect_2afc_exp2
   a) dicodetect_2afc_exp2 is an audio threshold experiment testing subjects' ability to determine the gender of a talker while other stimuli are presented.
   b) Subjects ran the experiment in an isolation booth while wearing Gen4 Access custom ear plugs.
   c) 8 subjects were tasked to run 25 blocks with a varying number of trials.
   d) Study began on November 27.
   e) Subjects are on target to complete dicodetect_2afc_exp2 mid-December.

3) Dicodetect_2afc_exp3
   a) dicodetect_2afc_exp2 is an audio threshold experiment testing subjects' ability to determine the orientation of a talker's voice (i.e. forwards or backwards) while other stimuli are presented.
   b) Subjects ran the experiment in an isolation booth while wearing Gen4 Access custom ear plugs.
   c) 8 subjects were tasked to run 15 blocks with a varying number of trials.
   d) Study began on November 27.
   e) Subjects are on target to complete dicodetect_2aft_exp3 mid-December.

4) deviceLocalization
   a) deviceLocalization is an audio localization experiment testing subjects' ability to localize ¼ second and continuous sound clips while wearing a variety of hearing devices.
5) Devices being tested are Combat Arms Ear Plugs (CAEP), Mini DTAC custom plugs, Nacre foam plugs, Peltor Comtac headset, Peltor Sportac headset, SensiMetric custom plugs, and Silynx foam plugs.
6) 9 subjects were tasked to run 16 blocks (2 blocks for each device plus 2 blocks with an open ear condition). Each block contains 180 trials.
7) Subjects are on target to complete deviceLocalization.
8) Nine subjects were scheduled daily to participate in the device Localization study. Earplugs and earplug systems tested were: Combat Arms Ear Plugs (CAE), Mini SensiMetric custom plugs, Silynx foam plugs, Bose headset, CEPS foam plugs, MSA Srdn headset, and SoComX custom plugs.
9) Six subjects were scheduled daily to participate in the missingSource study.
10) Six subjects were scheduled daily to participate in the missingSource, missing Source2 and missingSource3 studies.
11) The Relevant Set Size in a Multiple Source Sound Localization Task study is approximately 60 percent completed.
12) Thirty six training blocks and 7 data blocks have been completed for the SOCOM Localization Study.
13) Data collection for the deviceLocalization study was completed.
14) Data collection is underway for the missingSource study, and will be completed during the next reporting period.
15) MissingSource and missingSource2 studies have been completed. Data collection for the missingSource3 study is currently underway, and will be completed during the next reporting period.
16) missingSource3
    a) Same setup as missingSource and missingSource2 with a few exceptions.

b) Subject exposure to the sound sets varies between trials at intervals of 1.5, 2.5, 4.5, 6.5, 8.5, and 12.5 seconds.
c) Subjects are exposed to 2 to 15 environmental sounds per trial.
d) Each block is designated as either an "onset" block or an "offset" block. For an offset block, subjects must localize the sound that disappears for the sound set (as described for missingSource and missingSource2). For an onset block, subjects must localize the sound that appears in the set after the sounds have begun.
e) Subjects must also verbally identify the sound they were localizing to the ALF controller. The ALF controller then enters the subject's selection into the program.
f) Six subjects were selected to complete twenty-four blocks of the experiment. Each block contains thirty trials.

17) missingSource4
    a) Same setup as missingSource3 with a few exceptions.
    b) All "onset" blocks contain 15 environmental sounds per trial. All "offset" blocks contain 6 environmental sounds per trial.
    c) Six subjects were selected to complete twelve blocks of this experiment. Each block contains thirty trials.
18) Six subjects were scheduled daily to participate in the missingSource3 and missing Source4 studies.
19) Subjects completed missingSource3 and missingSource4.
20) Replaced bad speakers in the Auditory Localization Facility (ALF) Chamber.
21) Replaced bad speakers and sent in the bad ones to the re-cone repair facility for the MIRE chamber
22) Replaced and repaired broken BNC cables in the REAT facility
23) The Audio/Visual Conjunction Search study has been completed.
24) Subjects completed missingSource3 and missingSource4.
25) Six subjects were scheduled to participate in the VisualFeatureSearch and VisualFeatureSearch2 studies.
26) The Relevant Set Size in a Multiple Source Sound Localization Task study is approximately 60 percent completed.
27) Thirty six training blocks and 7 data blocks have been completed for the SOCOM Localization Study.
28) 11 subjects were scheduled daily to participate in the SOCOM Localization Study.
29) Eight subjects were scheduled daily to participate in the A/V Conjunction Search Study.
30) The SOCOM Localization Study has been completed.
31) The A/V Conjunction Search Study was completed
32) Eight subjects were scheduled daily to participate in the Bandwidth Effects in Multisource Localization Study.
33) The Bandwidth Effects in Multisource Localization Study has been completed.
34) The Bandwidth Effects in Multisource Localization Studies 3, 4 and training for 5 have been completed.
35) The SOCOM PPELOC study has been completed.

REAT Facility

1) 1342 Gentex HGU-84/P TPL liner, visor down, STD earcup w/speaker; 1343 Aero Combat Arms w/acoustic switch (pointed toward ear); 1344 Aero Combat Arms w/acoustic switch (pointed away from ear)
2) Thirteen attenuation studies were completed during this reporting period:
    a) 1304 – Gentex HGU-56P + ACCES Gen 5 Aircrew earplugs, 1306 - Gentex HGU-56P + ACCES Gen 5 Groundcrew earplugs, 1310 - Acousticom 5838-CA + ACCES Gen 5 Groundcrew Earplugs, 1319 – Gentex HGU-55P + ACCES Gen 5 Aircrew earplugs, 1320 – HGU-25P w/Goggle and Safety Direct Aural Protector + ACCES Gen 5 Groundcrew earplugs, 1346 – Gentex HGU-84/ Clear Visor, TPL Liner Std. Cups with speaker + CEP505-C11 Hearing Components Comply Canal Tips, 1347 – Red Tail Hawk Modified Custom Earshell, 1348 – Red Tail Hawk Modified Headset + Red Tail Hawk Modified Custom Earshell, 1350 – David Clark 40493G-01 w/ACCES cable mod Oregon Aero

83006DM Hush Kit, 1351 – MSA ACH 2002 + Oakley SI M Frame 2.0Z87 + Peltor Comtac, 1352 – MSA ACH 2002 + Oakley SI M Frame 2.0Z87 + Silynx Quiet Ops (fit check, passive test) and Hearing Components Comply Canal Tips, 1353 – Oakley SI M Frame 2.0Z87 + Peltor Comtac (passive test), 1355 – David Clark 40493G-01, Oregon Aero SoftTop Headset Cushion, ACCES cable mod 10405G-05..

b) REAT studies 1317, 1318, 1354, 1356 and 1363 were completed during this reporting period.

3D Facility: Acoustic subject panel availability and overall operation was scheduled and monitored for the following studies:

1) **Aswitch_whisper:** The study evaluates target intelligibility with noise-vocoded as well as normal speech targets as a function of signal-to-noise ratios when both were interleaved in time.
2) **MultiMRT:** The study assesses intelligibility of two MRT phrases when both were time-compressed at the different rates and presented with or without a delay to each ear.
3) **MRT_GH:** The study measures target intelligibility as a function of signal-to-noise ratio when the signal was processed through 3 different types of vocoders.
4) **MRT_GH_Tandem:** The study measures target intelligibility as a function of signal-to-noise ratio when the signals were processed through 2 different types of vocoders placed in tandem.
5) **Darpa Classify:** The aim of this experiment is to measure the ability of listeners to classify up to 5 different helicopter sounds in quiet when they hear a 1 sec sample of recorded flyover.
6) **DarpaDetect2:** This study evaluates the listener's ability to detect and identify a 1 sec snippet of a helicopter flyover when presented along with 6 different ambient sounds.
7) **MultiMRT3, MultiMRT4 and MultiMRT6:** These studies assess the listener's ability to report an MRT word heard in each ear as the phrase lengths of the words and the overlap between the words in the two ears were varied.
8) **Alfbandwidth:** This study examines the importance of high-frequency information for auditory localization in the presence of a masker.
9) **MultiMRT3, MultiMRT4 and MultiMRT6:** These studies assess the listener's ability to report an MRT word heard in each ear as the phrase lengths of the words and the overlap between the words in the two ears were varied.
10) **MRTAsynch3 and MRTAsynch4:** This study determines whether the tolerance for audio-visual asynchrony depends on the speaking rate of the talker.
11) **DarpaDetect_add:** This study evaluates the listener's ability to detect and identify a 1 sec snippet of a helicopter flyover when presented along with 6 different ambient sounds.
12) **Darpadetectbinaural:** This study was conducted in order to assess the additional benefit provided by two ears as opposed to a monaural listening condition.
13) **MultiMRT3, MultiMRT4 and MultiMRT6:** These studies assess the listener's ability to report an MRT word heard in each ear as the phrase lengths of the words and the overlap between the words in the two ears were varied.
14) **MRTAsynch3 and MRTAsynch4 and MRTAsynch6:** These studies determine whether the tolerance for audio-visual asynchrony depends on the speaking rate of the talker.
15) **Binaural_sparse2:** This experiment evaluates listeners' ability to understand spectrally and temporally sparse signals when binaural cues were presented.
16) **Bugstream_morse, Bugstream_morse2, Bugstream_morse3:** These experiments are designed to investigate stream formation using speech stimuli when pitch characteristics and the rhythm change between the target and masking streams.
17) **DarpaDetect2_add:** This study evaluates the listener's ability to detect and identify a 1 sec snippet of a helicopter flyover when presented along with 6 different ambient sounds
18) **MRTAsynch6:** This study determines whether the tolerance for audio-visual asynchrony depends on the speaking rate of the talker

MultiSound Local: Facility controller for ALF (Audio Localization Facility). Ran subjects in the experiments listed below:

1) missingSource
   a) missingSource is an audio localization experiment testing subjects' ability to localize looped environmental sound clips (dogs barking, pigs squealing, soda pouring, etc.).
   b) The four conditions of the experiment pertain to the length of time the subject is exposed to each sound set. The conditions are 2.5 seconds, 4.5 seconds, 6.5 seconds, and 8.5 seconds.
   c) For each trial, subjects are exposed to a variable number of environmental sounds (2 to 12).
   d) The sound sets begin simultaneously. After the designated length of time, (2.5, 4.5, 6.5 or 8.5 seconds) one sound will disappear from the set. The subject must then localize where the missing sound was.
   e) Sounds only appear along the horizontal plane.
   f) Six subjects were selected to complete eight blocks of the experiment. Each block contains forty trials.
2) missingSource2
   a) Same setup as missingSource expect the length of time subjects are exposed to the sound sets varies from trial to trial within the blocks.
   b) Six subjects were selected to complete six blocks of the experiment. Each block contains thirty trials.
3) missingSource3
   a) Same setup as missingSource and missingSource2 with a few exceptions.
   b) Subject exposure to the sound sets varies between trials at intervals of 1.5, 2.5, 4.5, 6.5, 8.5, and 12.5 seconds.
   c) Subjects are exposed to 2 to 15 environmental sounds per trial.
   d) Each block is designated as either an "onset" block or an "offset" block. For an offset block, subjects must localize the sound that disappears for the sound set (as described above). For an onset block, subjects must localize the sound that appears in the set after the sounds have begun.
   e) Subjects must also verbally identify the sound they were localizing to the ALF controller. The ALF controller then enters their selection into the program.
   f) Six subjects were selected to complete twenty-four blocks of the experiment. Each block contains thirty trials.
4) Subjects completed missingSource3 and missingSource4.
5) Replaced bad speakers in the Auditory Localization Facility (ALF) Chamber.
6) Replaced bad speakers and sent in the bad ones to the re-cone repair facility for the MIRE chamber
7) Replaced and repaired broken BNC cables in the REAT facility
8) ALF facility controller for BandwidthStudy and BandwidthStudy2.
   a) *BandwidthStudy* is an audio localization experiment testing a subjects' ability to localize a variety 'clicks' with discreet bandwidths.
   b) *BandwidthStudy2* is a follow-on study with a different set of 'clicks' with discreet bandwidths.
9) Helped calibrate ALF.
10) **10)** BandwidthStudy and BandwidthStudy2 complete.
11) ALF facility controller for BandwidthStudy3, BandwidthStudy4, and BandwidthStudy5.
12) *BandwidthStudy3* is an audio localization experiment testing a subjects' ability to localize a variety 'clicks' with discreet bandwidths.
13) *BandwidthStudy4* is a follow-on study of BandwidthStudy3 with a different set of 'clicks' with discreet bandwidths.
14) *BandwidthStudy5* is an audio localization experiment testing a subjects' ability to localize a variety 'clicks' with discreet bandwidths in the presence of a masking sound.
15) BandwidthStudy3, BandwidthStudy4 and BandwidthStudy5 are complete.

VOCRES Facility

1) Ten subject panel members were scheduled to participate in MRT data collection for Active Extreme.
2) Subjects were recruited and physical measurements of seven subject panel members and three ad hoc subjects were taken, in anticipation of the evaluation of the intelligibility of the upgraded intercom system of the NASA space suit.
3) Ten subject panel members were scheduled to participate in data collection for the GenVEAR study
4) Data collection for the GenVEAR study is completed.
5) One subject panel member and two ad hoc subjects were scheduled for training in VOCRES, in anticipation of the evaluation of the intelligibility of the upgraded intercom system of the NASA space suit.
6) Six Summer Panel members are participating in the SOCOM Speech Intelligibility Study in VOCRES.
7) Training was completed.
8) Six Summer Panel members were scheduled to participate in the SOCOM Speech Intelligibility Study in VOCRES.
9) Six panel members were scheduled to participate in the NASA Communication System Intelligibility Study in VOCRES.
10) The SOCOM Speech Intelligibility Study has been completed.
11) The NASA Communication System Intelligibility Study has been completed.
12) **ITDLevel2/ILDnew2:** This experiment uses spectro-temporally sparse signals in order to evaluate the benefit provided by binaural cues, such as interaural time and level differences, in target segregation tasks.
13) **Talk3_oddmasker:** This experiment evaluates the decrease in performance when the third masker is added as a function of the masker characteristics.
14) **Binaural_sparse:** This experiment assesses the role of both binaural cues (ILD/ITD) in understanding a target talker, when the talker is spectrally sparse.
15) **Dichomask2:** This experiment evaluates intelligibility when a masker in the contralateral ear and its temporal characteristics is varied.
16) **Stanford Gregory Kent:** This experiment assesses the feasibility of using an enhanced display to increase time of completion of a given communication task.
17) **Binaural_sparse2:** The experiment evaluates listeners' ability to understand spectrally and temporally sparse signals when binaural cues were presented.
18) **Dichomask4:** This experiment assesses the effect of a dichotic masker on segregation performance in the ipsilateral ear as a function of SNR and type of masker.
19) **Mask_Allocate2:** This experiment examines the performance in four different dichotic target segregation tasks as a function of type of masker when performance has been equated in the diotic condition.
20) **TMMOVE2:** The experiment assesses listeners' ability to track a moving target while the masker locations were fixed as a function of speed and extent of target movement.
21) **TMMOVE3:** The experiment assesses listeners' ability to track a moving target while the masker locations were fixed as a function of speed and extent of target movement.
22) **Talk3_oddmasker3:** The experiment investigates the role of the call sign in multimasker penalty.
23) Eleven acoustic subject panel members were scheduled to participate in the Evaluation of the Navy Noise Cancelling Microphone in VOCRES.
24) The Evaluation of the Navy Noise Cancelling Microphone has been completed.

CAVE Facility

1) The Egocentric Cueing of Exocentric Information in Urban Operations study has been completed.

Radio Laboratory

1) The Vocoder Intelligibility study completed.

Effect of Variable Visual Feedback Delay on Target Acquisition Performance
1) The Effect of Variable Visual Feedback Delay on Target Acquisition Performance Study completed.

Acoustic subject panel availability and overall operation scheduled and monitored for the following studies:
1) **Mask_Allocate2:** This experiment examines the performance in four different dichotic target segregation tasks as a function of type of masker when performance has been equated in the diotic condition.
2) **MultiCRM Studies (1, 2 and 3):** These studies evaluate the efficacy of presenting time-compressed stimuli consecutively instead of simultaneously for the purposes of designing an optimal auditory display.
3) **Aswitch_whisper:** The study evaluates target intelligibility with noise-vocoded as well as normal speech targets as a function of signal-to-noise ratios when both were interleaved in time.
4) **MultiMRT:** The study assesses intelligibility of two MRT phrases when both were time-compressed at the different rates and presented with or without a delay to each ear.
5) **MRT_GH:** The study measures target intelligibility as a function of signal-to-noise ratio when the signal was processed through 3 different types of vocoders.
6) **MRT_GH_Tandem:** The study measures target intelligibility as a function of signal-to-noise ratio when the signals were processed through 2 different types of vocoders placed in tandem.

Interactive Team Dialogue Effectiveness Evaluator (ITDEE) Study
1) In an effort to evaluate which of five communication environments enable four individuals working together as a team to communicate effectively, four teams of four subject panel members participated in the ITDEE study that was conducted in VOCRES. The ITDEE task is a type of communicability test that looks at both communication effectiveness and the communication environment. The five communication environments that were evaluated included looking face to face, communicating while separated by cubicles, conference call, VOIP, and chat.

Earplug Material and Construction Study
1) Data collection has been completed in REAT for one panel member using nine different types of earplugs. This study is designed to determine the effect of different earplug materials and construction on attenuation.

Acousticom Active Noise Reduction (ANR) Study
1) Data collection has begun for ten subject panel members for the Acousticom ANR study. Each subject was fitted with a Gentex helmet and ANR ear cups and tested in REAT and MIRE. This study is designed to test the attenuation of the Gentex helmet and Acousticom ANR ear cup combination.

Gen4 ACCES Plug Study
1) Data were collected from nine panel subjects in the REAT facility to determine the attenuation of the Gen 4 vented Acces plugs when worn with a 56P helmet.
2) Earmolds have been made for the four subject panel members who will be participating in the study.
3) Data collection completed**.**

1233 Gentex HGU-56P with David Clark ANR earcups minus custom plugs
1) Seven subject panel members participated in a study in the MIRE facility in which the attenuation of Gentex HGU-56P with David Clark ANR earcups was measured with the Acces Gen 4 aircrew earplugs.
2) Data collection completed.

Adaptive Technologies (ATI) Earmolds

1) Seven panel subjects had earmolds made in support of the ATI study.

Combat Search and Rescue (CSAR) Study

1) In an effort to examine the impact of auditory cueing on finding a visual target in the context of a simulated Combat Search and Rescue task, six subjects completed two training runs each for the CSAR study in the CAVE facility. The subjects were tasked with navigating through a "maze" which appears as a forested region or an urban scene and find a target (simulated human) with the aid of a directional cue (clock coordinate displayed visually) or a spatialized auditory cue that comes from the direction of the target.
2) Data were collected for five panel subject members participating in the CSAR study in the CAVE facility.

FITTS Study

1) Six subject panel members are scheduled to participate in this study, which is designed to study the effect of variability of delay on motor performance. Two subjects have begun training.

TAC Auditory Localization Study

1) Six subject panel members have completed 75 sessions in support of an effort to test localization with TAC plugs using various algorithms.
2) One of the subject panel members was trained to run the facility and is now assisting with data collection.

FAA Training Program

1) Four subject panel members participated in testing an FAA training program, which will be used to train air traffic controllers to do critical listening.
2) Data collection completed.

1281 Solid Vinyl Plugs with Creare STTR

1) Three ad hoc subjects were scheduled to participate in a study in the REAT facility in which the attenuation of Solid Vinyl PVC plugs worn with Creare STTR was measured. Data were successfully collected.
2) Data collection complete.

1286 JSF ANR ACCES Plugs

1) Two subject panel members and three ad hoc subjects were scheduled to participate in a study in the REAT facility, in which the attenuation of JSF ANR Acces custom plugs was measured.
2) Two subject panel members participated in a study in the REAT facility, in which the attenuation of JSF ANR Acces custom plugs was measured.
3) Data collection complete.

1287 Gentex HGU-56P with David Clark ANR earplus (passive) using Acces Gen 4 Aircrew Plugs

1) Nine subject panel members and one ad hoc subject were scheduled to participate in a study in the REAT facility, in which the attenuation of Gentex HGU-56P with David Clark ANR earplugs (passive) using Acces Gen 4 Aircrew Plugs was measured.
2) Nine subject panel members participated in a study in the REAT facility, in which the attenuation of Gentex HGU-56P with David Clark ANR earplugs (passive) using Acces Gen 4 Aircrew Plugs was measured.
3) Data collection complete.

1288 JSF ANR Acces (unpopulated) with Aegisound Max 25/40 Earmuff

1) Seven subject panel members and two ad hoc subjects were scheduled to participate in a study in the REAT facility, in which the attenuation of JSF ANR Acces (unpopulated) with Aegisound Max 25/40 Earmuffs was measured.

Environmental Sounds Study

1) Preliminary pilot data have been collected from ten subject panel members for the Environmental Sounds Study. The subject's task is to match a word with an environmental sound, in an effort to determine if there are meaningful words that can be used as warning signals, instead of sounds.

HGU-56P with Sound Guard Earplugs Study

1) Data collected from three subject panel members in the REAT facility to determine the attenuation of Sound Guard earplugs when worn with a 56-P helmet.
2) Data collection complete.

Gen 4 Acces ANR with 55P Helmet Study

1) Data were collected from three subject panel members in the REAT facility to assess the attenuation of Gen 4 Acces ANR with 55 P helmet was studied.

2) This study is on hold since the ear cups had to be returned to the company prior to completion of the study.

CueingExp

1) Open ear audio localization experiment testing subjects' ability to localize ¼ second environmental sound clips and/or sound clips of spoken phonetically balanced (PB) words.
2) For each trial, subjects were given either a pre-cue or post cue clip of what sound or word they would be localizing.
3) For each trial, the environmental sound clip or PB word was presented in either forward or reverse motion.
4) For each trial, the environmental sound clip or PB word was presented with 0 to 5 masking sounds or words respectively.
5) 8 subjects were tasked with 56 blocks of ¼ second sound clips. (* 1 block = 50 trials)

3D Global Hawk Communication Study

Project Status Summary: The test objective is to measure the effects of continuous variable slope delta (CVSD) and in tandem with CVSD, adaptive differential pulse code modulation (ADPCM), and voice over internet protocol (VoIP) vocoding algorithms on speech intelligibility over an ARC-210 radio link. These components are considered the critical links in the air traffic controller (ATC) to global hawk mission control element (MCE) communication path. Generally, the guidelines in ANSI S3.2-1989, measuring the intelligibility of voice communication systems, will be followed by using the modified rhyme test (MRT). Speech intelligibility will be evaluated with the ARC-210 radios in non-secure, secure, and HAVE QUICK II modes in the simulated communications path between the ATC and MCE stations and in real communications via an INMARSAT communications link. The communication system must achieve a required mean intelligibility level of 80 percent with the MRT to be considered acceptable by the operator.

Intelligibility Measurements of CVSD, ADPCM and VoIP Vocoding Techniques:

1) Identification and configuration of a PC development system for support of the vocoding techniques implementation.
2) Exploration and study of the Natural Access Open Source software libraries and its application for the AG 2000-BRI platform PC card for telecommunications systems.
3) Assessment of which service functions are applicable to the goals of the project and what support services are required and wrap them into a telecommunication services class for use in Windows application development. This work is currently on-going and is in the early stages and is expected to carry forward into the next period.
4) Work to port the (DDVPC) DoD Digital Voice Processor Consortium MELP algorithm to the TI C5510 platform continued this period. Work consisted this period of conforming all dynamic

memory allocations over to the DSP/BIOS environment as memory allocation outside of DSP/BIOS management would fail. The existing memory allocation macros were converted to DSP/BIOS memory allocation calls. The new method required the definition of a memory heap in the DSP/BIOS (MEM) Memory Object, the address of which was then referenced in the allocation calls. To complete the adaptation in DSP/BIOS the input and output data streams were converted to (PIP) pipe objects in two stages: one for the encoder and one for the decoder. (SWI) Software Interrupt modules were defined to manage the processing of the PIP buffers and schedule the data processing handlers which effected the encoding and decoding processes. This SWI and PIP structure was previously checked out with identical length buffers at all stages for pass through of audio data. The buffer lengths of the receive and transmit pipes were then matched to the frame lengths of the original algorithm, 180 words. The intermediate pipe buffer lengths were matched to the encoded channel lengths of six words each. In order to enhance the debug and trace capability of the algorithm, all of the (printf) formatted print calls were converted to a DSP/BIOS (LOG) log object for display of error messages and other trace information. This adapted algorithm structure was made to compile and build to an executable file which was able to be loaded to the DSP platform by the TO Code Composer Studio IDE debugger. Running the loaded executable produced a data sorting error being reported to the LOG object and the algorithm execution would abort. It has been decided not to spend effort to investigate the cause of the sorting error, but rather to wait for the receipt of the optimized for C5510 platform MELP encoder/decoder object from Adaptive Digital Technologies. The knowledge gained in applying the DDVPC algorithm to the DSP/BIOS environment will be applicable at that time.

5) The MELP algorithm library objects were received from ADT and implemented in a demonstration project on the TI C5510 DSK board. A first attempt was developed that utilized software pipes and software interrupts to manage the propagation of the receive, encode, decode and transmit functions. The first implementation proved unsatisfactory because it did not support programming of the DSK codec sample rate for the prerequisite 8 kHz. This first attempt also performed manual sorting of the stereo channels to internal buffers and did not take advantage of the (DMA) Direct Memory Access capability to sort the stereo channels for segregated processing. A suitable software template was identified and used for the more successful demonstration. This project made use of hardware interrupts to schedule software interrupt modules to perform the various functions; pipes were not used, instead "ping" and "pong" buffers defined in on chip memory were utilized. This configuration allowed for the programming of the codec sample rate for the required 8 kHz and also incorporated the DMA function and employed the channel sorting function of the DMA. Two demonstrations were developed. Both split up the left and right channel processing in separate threads; however one performs encode and decode functions in a single module and the other separates the decode and encode functions into separate modules. The later configuration was developed with an eye for the future need to separate encode and decode functions on separate transmit and receive platforms. Timely and insightful help from the ADT technical staff is acknowledged as being instrumental in achieving this milestone.

<u>Soft Phone (VoIP) Data Stream Rendering:</u>

1) Work to build and evaluate the JVOIPLIB VoIP algorithms is being transferred to another task which overlaps VoIP requirements.

2) Access to the 2005 MS Visual Studio C++ was arranged for building the JVOIPLIB test utility; however, the application failed to build under that development environment. A file present in the project collection was found to permit conversion to the 2003 development environment. When employed, and with a TYPEDEF definition for a unsigned integer data type, the

application build completed successfully. The libraries were then built with the 2003 version. All modules were installed to three Network Evaluation Facility computers. The computers were networked via a network switch. VoIP sessions were established using the test utility application. A subjective evaluation of JVOIPLIB indicates that the library should provide a basis for integrating VoIP into (SLAB) NASA Ames Sound Lab as a sound source and HECB will discuss the effort with the NASA Ames developer. This will preclude the need to implement soft phone/VoIP capability in the lab using soft phone open source software like sipXphone or sipXezPhone.

3) Facility computer resources have been checked out for operability. One computer failed to boot and another was found to have a bad graphics interface. The one which failed to boot has been turned in; a request was submitted to the TM via email as to how to proceed with the bad graphics interface. Currently there are five working PCs and two archival machines.

4) A meeting was held with the Task Monitor, and other branch personnel to further identify requirements to the subcontractor, VRSonic, for the VoIP sound source capability for SLAB. Topics discussed included a receive VoIP sound source associated with a unique IP address, a rewind buffer for replay capability with voice detection to accomplish “catch-up”, and channel power statistic query. The previously evaluated JVOIPLIB VoIP software will be proposed to the sub-contractor as a possible source or model for the VoIP processing. The results of the meeting are to be presented to the contractor during a face to face meeting set for the following week by branch personnel.

5) No additional work was accomplished on facility hardware setup and configuration due to lack of further direction from the TM.

3D Audio Chamber Studies: Subject panel availability and overall operation was monitored for the following studies:

1) Detect Tone and Noise studies which validate thresholds.
2) Whisper which evaluates target intelligibility with multiple whispering talkers, in order to assess target segregation efficacy in situations where talkers are required to be unobtrusive.
3) Bands_grouping which assesses the ability of listeners to identify a target signal under 3 experimental conditions: 1) When the target and masker had unique fundamental frequencies, 2) When the target and masker shared the same fundamental frequency, and 3) when the target contained some of the fundamental frequency information of the masker and vice versa.
4) Grouping_control which assesses if the presence of a call sign aided target identification with artificial speech signals, where segregation was found to be difficult.
5) SpeedCP which assesses the influence of rate of speech on target segregation in a multitalker listening task.

Anechoic Lab

1) Four subject panel members were scheduled to participate in a Headphone Repeatability Study (EqualizationFrameSW2) in the Anechoic Lab.
2) Data collection has been completed for three of four subjects for this study.
3) Three subject panel members were scheduled to participate in a Headphone Repeatability Study (EqualizationFrameSW2) in the Anechoic Lab.
4) One subject panel member was scheduled to participate in a Headphone Repeatability Study (EqualizationFrameSW2) in the Anechoic Lab.
5) Data collection for this study has been completed.

USSOCOM

1) Rewire a microphone adapter / power box for data collection on the upcoming C-17 aircraft in-flight data collection
2) Completed and tracked several purchase requests and orders of equipment and materials for this program

3) Tasked and had EMI testing completed through the Sensors Division on a National Instruments system to collect in-flight data on the C-17 aircraft at Charleston, AFB SC, TDY and flight tentatively scheduled for 15 – 19 Oct 2007
4) Started packing equipment and performing preliminary tests on data collection system
5) Traveled to Charleston Air Force Base in South Carolina to take in-flight C-17 recordings.
   a) Took measurements of the cargo area and cockpit to retrofit cabling for recording system.
   b) Cut cables to length for hook up inside plane.
   c) Ran cables, hooked up microphones, and calibrated recording system.
   d) Took notes while in-flight maneuvers were recorded.
   e) Dismantled recording system and paced it up for shipment back to Dayton.
6) Ordered and modified several David Clark Headsets and Pilots Helmets for the C-17 Aircraft at Charleston, AFB SC
7) Traveled to Charleston and collected in-light noise data in the C-17 Aircraft
8) Completed and tracked several purchase requests and orders of equipment and materials for this program
9) Start running the OHMS study.
10) Start preparing to run another Cornell/RoboFlag study
11) Facility controller for VOCRES (Voice Communication Research Evaluation System). Trained five subjects on the basic functions of the facility. Explained the purpose and logic behind the MRT Test (Modified Rhyme Task) and demonstrated how to use the touchpad computers and user interfaces to complete the task.
12) Subjects were trained in 'quiet' and 'pink noise' environments.
13) Subjects 45, 47, 1349, 1365, and 1379 were trained during this session.
14) Subjects are fully trained in VOCRES

Net Centric Communications Support: Continued progress is reported for the (MMC) Multi-Modal Communication Monitor software program. Additional features have been added including speech activity, recognition of speech content for brevity formats display, and incorporation of a simple keyword search capability. Three demonstrations of the system have been conducted for soliciting user community assessment and criticism.

Enhanced MMC Monitor Software Development:

1) The (IPSS) Internet Protocol Server SLAB Server was updated to support SLAB version 6.1.0 functionality. This supports the capability to mix allocated DIS Radio sources based on frequency and render them as a single source. This has been shown to work reliably making the allocation based on channel ID to be obsolete. This capability needs to be stripped from the GUI so that mapping of the DIS source is tied to frequency selection instead of channel ID. Commands have been added to the IPSS to support the allocation of DIS radio sources and for the mixing of DIS radio sources tuned to a given frequency. Other commands that provide DIS radio statistics and information were altered in the latest SLAB version and have been modified as appropriate. Notes pertaining to the new features are being maintained for updating the IPSS User Reference.
2) Work to effect the capturing of analog sources from an ASIO streaming device is progressing. The DIS transmit software has been altered to allow the command line selection of the desired capture device number, so that the Windows default device specification can be left as is. Also, a capture data routine was written to parallel the original capture routine, but to also down sample the captured data by an integral divisor of the captured sample rate. Capturing at 48 kHz the data can be simply down sampled to 8 kHz which is readily supported by the present WCAS software. The capture algorithm also supports channel level activity detection and enables / disables the transmitted as needed (akin to hot mic or VOX).
3) The related work effort to effect reliable frequency tuning is still in progress. The MMC monitor program features were added to query the local machine (or any named machine on the network) IP address and use this information to set the DISNET address and to use the last portion of the IP to define a unique DIS network application number. This together with an MMC application number (defined by the assigned IPSS sound source number) used as

the DIS entity number defines a unique DIS net channel ID for each MMC application on the network. This also facilitates the programmability of the wave file location for MMC non-real time playback support. The DIS transmit frequency is given by a parameter in the DISSound.ini file. While this seems to work well for DIS frequency tuning for SLAB, the transcription of the transmitted phrase does not show up in the appropriate tuned MMC Monitor application.

4) The (MMC) Multi-Modal Chant Monitor software was readied for demonstration at the CORONA Tops meeting this month. Additional timing control was built into the startup procedure and command line parameters were added to support initialization of sound position; the timing control insured that the monitor application windows started in a predictable order that presented a uniform demonstration across all work stations. Over a two day period the CORONA Tops meeting was prepared and supported; the demonstration was successfully presented one-on-one to more than 15 generals and VIPs.
5) Following the CORONA the multiple-selection of transcribed utterances capability, the technology (based on the Open Source PortAudio cross-platform audio API) for which was developed earlier but too close to the CORONA for reliable integration, was added to the MMC Monitor software. Exercising of this new feature "on the bench" indicated a good new enhancement for the MMC Monitor (UI) user interface.
6) The suggested fix for the misalignment of utterances by frequency in WCAS during heavy DIS traffic was implemented. The fix is to use the most recent version of the waveExtract.exe program and effect a one line of code change in the PDUSorter.cpp file within the slabDISInterface library. Testing indicated that the misidentification was persisting. A work-around was developed that involves the generation of a properties file that equates call sign with channel ID and frequency. The frequency from this property file is used to override the WCAS-stipulated frequency and display the text in the proper window. If there is no entry in the property file for a given entity, then the WCAS frequency is used.
7) A meeting of the MMC project group was attended to support and plan the MMC tool analysis study and data collection process. Use of the analog data streaming to DIS network transmitter, the prototype of which was developed earlier, will be used to transcribe random utterances within the MMC speech-to-text domain. Once captured as wave files they can be transmitted over a DIS network displaying other traffic. The study will seek to measure how well subjects respond to certain commands presented in this manner. Support for utilization of the MMC monitor software for the generation of the command utterances and distracter phrases was provided as needed.
8) An exploratory web application is being developed to investigate this platform as a means of providing an MMC Monitor browser debriefing tool. The browser application currently will parse the XML results file generated by the MMC Monitor software during the mission or training session, and present the transcribed utterances on a web page. Each utterance is displayed with a button which supports selection and playback of the associated wave file. The text display can also be made to support enhanced display properties such as font style, color and highlighting at the word level. Other editing or notation features could possibly be added by associating a context menu with the selection button. A problem with this format is that user interface controls positioned at the top of the page are lost as the text is displayed causing the web page to scroll controls out of view. A means of "floating" a control panel will need to be implemented. A project status meeting will need to review and determine the merit and utility of this effort.
9) The (MMC) Multi-Modal Chat Monitor software continues to have work done to add new features and usability. New features include text (transcribed utterance) tagging (including multiple selections) and maintenance of a local XML results file to support re-populating the text display from the beginning of the mission when the monitor frequency is changed. The use of the local results file also supports running the MMC tool in a "Debriefing mode". In this mode the MMC is run in stand-alone mode (without WCAS and SLAB servers). The program UI has an "Open file dialog" button to allow the user to select from XML results files from previous missions. Available frequencies appear in the frequency select drop-down control and selecting one populates the text list of comm. associated with that frequency. The user can scroll, search, select and play utterances the same as in the real time mode. A show

tagged (flagged) items command button has been added to the UI. Other enhancements include use of pictorial labels on select command buttons.

10) Follow-up demonstration and support for integration of the MMC software in the CTT lab in RHCP were performed. Some usability enhancements for the UI were noted and implemented. Work to integrate the MMC is started and will continue into next month. To date the MMC, running on a RHCB laptop, has been shown to work on the CTT network with WCAS hosted on a CTT platform. The next goal will be to install and run MMC on several CTT platforms.
11) Two MMC demonstrations were supported for RHCI personnel and a meeting was attended to explore the possibility of using MMC in a UAV control simulation workstation environments. This work exploration will continue into next month.
12) Work was performed to examine the possibility of developing a MMC browser based application to be used as a MMC debriefing tool. The browser based application proved not to be a satisfactory platform for the MMC layout. It was anticipated that HTML would support more text formatting options so that keywords as well as word confidences could be readily tagged. However, browser list box controls do not support object items or even multiline items. Tables were necessary to display multiline utterances which needed to be generated dynamically. Also a way to make them independently scrollable, or to float other controls on the page while the page is scrolled to view content, was an issue. While Ajax ASP.net controls have been developed to do these things, the dynamic generation of the table components and content along with the need to persist the content after “round trips” to the server every time a button control was clicked by the user proved too problematic. The development was abandoned in favor of the more robust MMC tool Debriefing mode described above. However, the desire to incorporate more rich text-like text formatting remains unresolved and a means to add Rich Text Box controls to display formatted text needs to be developed.
13) Enhancements for the (MMC) Multi-Modal Chat Monitor user interface were accomplished. It is now possible for users to make corrections to transcribed text and also to annotate utterances with comments. A context menu is available for performing these and other functions by right-clicking the list box area. Text tagging, re-playing utterances, selecting previously tagged, corrected or annotated transcriptions, raw (unformatted) text display, and auto-scroll enabling are also supported by the context menu items. A new mode of operation is also now supported that nullifies the spatial audio effect and disables the speech transcription display for the MMC usability study; this mode is enabled by setting the sound position parameter in the command line to zero.
14) MMC installs at RHCP (CTT lab) and RHCI are in progress. The RHCP install was attempted and the target platform was determined not to be fast or powerful enough to support the WCAS and SLAB and multiple MMC instances; RHCP is looking for alternative host machines. At RHCI the opposite was the problem. The high speed platforms there uncovered timing issues with the MMC applications interfacing to the IPSS manager client/server. To alleviate the timing concerns more handshaking is being built into the MMC application that will hold off the initialization of subsequent MMC instances until one instance is fully configured by the server. The first cut at the coding changes have been implemented and are being tested.
15) Two major accomplishments were completed for (MMCMP) Multi-modal Chat Monitor Plus application. The first is additional inter-instance handshaking among multiple instances of the MMCMP during initialization. All instances wait to be signaled before initializing their individual SLAB environments. The last instance (also the render control instance) will begin rendering and then signal the first instance to proceed and then wait to be signaled in turn. The first will finish its SLAB environment set-up and then signal the next in sequence. This continues until the render control instance is signaled. Complimentary code changes were made in the IPSS Manager client/server software to send signal command messages to the appropriate client instance when requested. In this way the running of multiple instances of the MMCMP hosted on fast multi-core processor machines will come up controlled and

orderly. This version of the MMCMP has been successfully installed on two workstations at RHCI.

16) Automation of DIS Log playing has been added to the MMCMP. The render control (last) MMCMP instance will look for a file that contains a list of DIS log files to play. If found the MMCMP will enumerate instances of DIS Log players running on the host system. If any are found the MMC will, in turn, send key commands to each player to load a DIS Log file from the list and generate a play button click message to begin playing it. It will do this for each file in the list provided that there is enough DIS Log player instances enumerated. This work was primarily accomplished in support of the MMC usability study.
17) In other MMC work, a need to speed up the WCas recognition times was discussed with the RHCP team. After talking with RHCP and demonstrating the problem, they decided on a course of action to dual-thread two recognizers in parallel. This solution has shown to virtually eliminate the long delays previously observed in the speech to text recognition. The solution also included an upgrade to a newer version of the WCas software which processed the results XML file differently. This new format did not work with the existing MMCMP software. Investigation of the cause showed that a node defining the entity ID was missing from the parent node. To correct the problem RHCP agreed to restore the missing node in the XML file. Although this new version is working fine on one workstation, attempts to port the new version to other workstations has been unsuccessful.
18) A MMC User Reference has been written as a Word 2007 document. It contains explanations with screen shots of the MMC Monitor Plus software (UI) user interface controls, command line parameters, usage modes, software dependencies and installation procedures. While an earlier version of this document previously existed (and may have previously been reported) this version has been updated to reflect the changes to the MMCMP of the past several months and expanded to include more complete installation and usage notes.
19) The current version of the MMC that supports DIS log automation were readied for deployment in the MMC usability study and for offsite demonstrations. For the former the updated software was installed on the experiment laptop and tested with the analog to DIS gateway software; in the latter a laptop workstation was prepared with the MMC software and WCAS running the ASKAS and JTAC speech models and briefed to the RHCB individual who was to perform offsite demonstrations of the integrated technologies. The latest version of WCAS (2007) is not yet installed for either of these platforms; an updated install of the WCAS 2007 is still pending from the WCAS developers to alleviate the porting issues observed with the newer version of WCAS.
20) The MMC User Reference document was delivered to the (ARL) Army Research Lab and also distributed to the research scientist conducting the usability study and the developer of the analog to DIS gateway.
21) Changes were made to the (MMC) Multi Modal Chat software to support a new version of the (WCAS) Warfighter Communication Assessment System (WCas 2007). A new installer for WCas 2007 was received and implemented on several of the MMC laptop workstations. The new installer alleviates the installation problems observed in the last report; however, other issues were encountered that required changes to the MMC program. The changes were made to accommodate new wave file folder specification and the inclusion of the sub-folder name in the wave filename which was affecting the parsing of the timestamp information. Although this is now working properly, other issues have been observed with the transcription accuracy. Consulting with the WCAS developers uncovered a configuration parameter set wrong that determined the grammar domain specification. Other issues needing to be resolved before deployment of the new software suite include proper sequencing of

transcription fragmentation and overlapping of fragments. Also a means of substituting different grammar model compilations for different environments is needed.

22) A capability to run the MMC with spatial audio but no text display was provided in support of the MMC Usability Study. This operation mode is controlled by making the spatial position parameter in the command line a negative number; the zero value setting of this parameter still controls the no text and no spatialization of the audio mode. Normal operation mode is effected by use of positive number values for the position parameter.
23) The Network lab is now established with the latest version of the WCas 2007 and the MMC Monitor software suite. All necessary configuration file modifications have been accomplished and the operation of the system checked out. A change to the MMC software was necessary to control the automation of the DIS Log Player start-ups on the lab PCs. The menu select key strike combinations beginning with the Alt-key needed to be separated into two command calls. In order for the system to work without memory access faults it was necessary to rebuild each of the PC systems without benefit of the standard AF desktop configuration. The memory faults were manifesting in the SLAB buffers and inhibiting the IPSS from running. Attention was given to transitioning the work to the designated government person and so this person was included in the work to build the network and install all of the software. Time was also given to a code overview of the MMC and IPSS software programs for this individual.

<u>General Aviation Flight Test Phase 2:</u>

1) In support of the next phase of GA 3-D audio flight tests, a program has been developed that will aid in the development and validation of the new flight test control program. The aid program displays the flight path of the waypoint task from the first phase experiments while transmitting the data from actual data collection files to the laptop computer running a modified version of the original flight test software. This version parses the data and extracts the ADAHRS flight data and generates the directional audio cue as the "simulated" plane advances through the waypoint course. Appropriate code changes have been determined and validated that uses the ADAHRS data to generate the sound cue position instead of the head tracker data, as was done as a work-around during the first phase of flight experiment data collection. During the next phase of the experiments, the head tracker will be utilized for subject head orientation measurements and cannot be used in the work-around method. The corrected code uses the flight angle (or azimuth track) reported by the ADAHRS together with the plane roll and pitch as the reference input to the vector transformation to generate the relative target location. Several of the first phase data files have been input to the "simulator" aid to validate the operation of the modified experiment control software. In each case the modified control program generates the appropriate sound cue using the ADAHRS data. To further validate the test procedure, the control program will be further modified to use the head data in a head relative mode and verify the operation with head-coupled sound generation. These validated algorithms will be incorporated into the second phase experiment control software yet to be designed. This work effort is on target for the anticipated check-out flights in early November 2008.
2) The (IPSS) Internet Protocol SLAB Server User Manual has been updated with commands and operation procedures now available in version 2.2.1 of the IPSS. These updates document the commands that perform DIS sound source allocation discriminated by frequency and the replay function.

3) The latest version of the IPSS software and documentation has been shipped on CD to the Army Research Laboratory HRED branch. An example program project was also included that exemplifies client program communication interface to the IPSS.
4) Data processing verification of ADAHRS data for both plane relative and head relative modes was completed this period. Additional archived data files from the first phase of the experiment were used to complete the validation. The previously developed "gaFlyer" software was used to read the data files and send them over RS-232 link to the laptop PC running the source positioning algorithms to be used for the Phase 2 experiment control. The validation consisted of listening to the navigation sound source position change as the "gaFlyer" graphically tracked the flight path of the plane during phase 1 runs. The positioning algorithms correctly generated the navigation sound source position for both head and plane relative modes. Work was then started on the design and development of the phase 2 GA Experiment control software. The (UI) user interface is currently being developed along with the file structure and I/O for the raw data file. The design plan is to develop the control software using C-sharp (C#) and use the IPSS for the control of the audio environment. This work will continue through the next period and is on target to be ready for use in early to mid November.
5) In support of ALF experiments changes were required for the (IPSS) Internet Protocol SLAB Server. A problem with allocating and playing wave files manifested. Investigation of the problem showed that setting a flag (a recent addition to the SLAB wave file allocation function) for SLAB to copy the wave file to memory before rendering, eliminated the problem. It was further determined that this problem may only exist for the first wave file being rendered, particularly if no other sound sources are being allocated. The IPSS was changed to support the memory copy flag and installed on the ALF computer. At the same time the Free-sources command, which previously did not work properly, was fixed so that it can now be used to free the SLAB environment without closing the SLAB instantiation.
6) The GA control and data collection software development progressed this period. The (UI) User Interface has had significant work accomplished as well as the communication sockets interface for the IPSS (largely adapted from the MMC software). Also worked is the wrapping of the Athena ADAHRS and IMU Head tracker interface software into a GA Support Library DLL. A C# version of the "ShowAthena" utility was developed to test the DLL interface. This has been shown to work with Athena data file captures previously provide by NASA during the 2006 development of the first phase software. A companion "ShowIMU" utility program has been written in C# so that the head tracker interface can easily be tested once on site at NASA. Other utility math functions, from the C++ "VectorAddition" class, have also been incorporated into the library for performing the axis transformations for rendering of the sound sources in 3-D space relative to the reference axis of the plane or the head tracker. Testing is currently being performed to see how well the IP interface to the SLAB server functions with the rapid updates required. If necessary the wrapping of the SLAB server into the GA Support Library will be investigated in order to gain greater speed in communicating updates to the server. Also, now all but finished, is the data manager class for writing and reading the experiment data for the Traffic Alerts portion of the experiment. Functions written include the constructor for opening or creating data files and functions for reading and writing the various data headers and records defined in the data model. To be done yet are the functions for generating the text data files for MATLAB analysis. A companion data file manager will be designed and written for the Displaced Attitude Recovery task. This will be simply a matter of adapted the existing file manager to the needs and requirements of the Recoveries data.
7) In support of GA software development changes were required for the (IPSS) Internet Protocol SLAB Server. A problem with setting the directory paths for the HRTF data sets

folder and the WAVE file folder was discovered and corrected; the "FreeSources" command also needed to restore the directory defaults. A command to reset SLAB from an error state was provided so that recovery from a missing wave file allocation could be effected; the link to a source command no longer calls SLAB Notify to avoid the unnecessary message box click-through.

8) Development of the experiment control and data collection software to provide needed functionality. The experiment control for the Recoveries scenario has now been coded; the Data Manager class has been expanded to include functions for reading and writing the Recoveries session data file. A Pause and Resume control was added to interrupt or suspend the Traffic Alerts data collection to allow for real traffic intrusions or other flight safety concerns. A special data record tag for specification of clock traffic alert onset and marking was provided to differentiate from spatial targets in the data file. Another data record tag indicates when a traffic alert suspension (canceling) occurs as a result of Pause mode activation. Improved control of the traffic mode picker keeps the desired percentage of non spatial traffic events within a tighter tolerance and forces the presentation of one or the other as needed as the data session proceeds. Volume slider controls for orientation cue, alert cues and overall volume are now provided and display labels for presenting plane and head attitude data are now part of the UI. Interaction of the various audio cues is now controlled so that a waypoint announcement will not coincide with a active clock traffic alert; similarly a new traffic alert onset will be held until a waypoint announcement is finished. Code has been added to secondary control event handlers to automatically return focus to the primary button control so that the Enter key can be used for marking an event (response).
9) A review of the UI was conducted with the Government Task Monitors and some additional requests were put forward. The activation of the orientation cue needs to be made present upon subject number validation and after the end of the data collection session, and a boresight function is needed to generate sound cues referenced to the orientation of the plane during straight and level flight. The Orientation cue has been made to be active upon validation of the subject number and before the start of the data collection session. The request to keep the cue active after the data collection is complete will be accomplished this period as will the implementation of the boresight reference.
10) More work was accomplished on the software program for the 3-D Audio General Aviation phase 2. Several new features have been added that now make the software as ready as possible prior to the integration testing and operational flight checks. Features added include a boresight function for horizon (orientation) level flight reference, maintaining of the horizon cue display after a data session, verification of the parsing of the INS data stream with a simulated ADAHRS data stream, incorporation of the IMU (head tracking) interface for parsing IMU data stream, plane orientation and waypoint distance information display, verification of reference translation for display of alerts in plane relative and head relative modes, and vigorous testing of the traffic alerts and recoveries paradigm and data collection.

AFRL/RHCI

Synthetic Interface Research for UAV Systems (SIRUS): The major projects are the Synthetic Vision 2 (SV-2) study, the Adaptive Levels of Automation 1 (ALOA-1) study, and the Vigilant Spirit 1 study. The Synthetic Vision 2 (VS-2) study examines user performance with various levels of synthetic vision overlay update rate for 4 realistic tasks in a UAV simulation environment. The Adaptive Levels of Automation 1 (ALOA-1) study examines user performance with several levels of auto-route automation for simultaneous supervisory control of 1,3, and 4 UAVs in a multi-UAV testbed. The Synthetic Vision 3 (SV-3) study examines user performance with 3 Picture-In-

Picture (PI) Levels and 2 Synthetic Vision Overlay Registration Error Levels (low, High) for large area search task in a UAV simulation environment. The Vigilant Spirit 1 study examines user performance with missions that require rapid task-switching in a multi-UAV testbed. The Predator Mapping Display project examines Predator operator issues with the TSD (Tactical Situation Display) and general human factors guidelines for development of a future TSD. AvantGuard project examines user performance with levels of automation for a supervisory task of 3 UAVs providing reconnaissance for a convoy through urban areas.

SV-2:

1) Completed SV-2 data collection (14 subjects total)
2) Organizing SV-2 data for analysis procedures
3) Generated statistical results and summaries as necessary or the SPIE 2006 paper, including target marking task data and update rate verification procedures.
4) Summarized synthetic vision overlay design guidelines based on literature review.

SV-3:

1) Analyzed data from 12 participants
2) Co-authored proposed paper for HFES (Human Factors and Ergonomics Society) 2007 annual meeting procedures
3) Worked on ideas for SV-4: control intensity of synthetic environment surrounding the camera view in PIP; vary the number of overlaid flags and synthetic elements; decluttering techniques; allowing real-time PIP level changes

ALOA-1:

1) Generated scenarios to demonstrate the likelihood of details of the ALOA-1 test plan.
2) Met with ORCA during their on-site visit
3) Worked on using Phase II SBIR software and providing feedback and software bugs to ORCA
4) Review final release of ALOA software and deliverables
5) Developed proposals for ALOA-2 study based on the latest software capabilities, namely levels of automation within the “Allocation Task”. Explored using Fidelity of Automation as an independent variable.
6) Developed sample scenario for the ALOA-2 design: reduced to 15 minutes and tweaked auto-routing and allocation parameters to reduce failures of automation.
7) Developed 7 trial scenarios and 2 training trials and tested them for the ALOA-2 study.
8) Generated, tested, and finalized experimental procedures, scenarios, and data logging processes.
9) Wrote Perl scripts to help analyze data.
10) Ran 2 “pilot” participants and then 2 full participants through data collection procedures.
11) Completed statistical analysis of ALOA-2 objective and subjective data.
12) Presented preliminary results.
13) Began writing Methods section for study publication.

Vigilant Spirit-1

1) Reviewed literature for task switching technologies and general information
2) Edited design document and software requirements
3) Brainstormed interface technologies and display concepts for consideration for multi-UAV support

4) Developed the CRM panel (Coordinate Response Measure) and HSD panel (Health and Status Display) software “plug-in” tools for the SVCS (Vigilant Spirit Control Station). The DCRM panel and HSD panel are designed to be secondary tasks for use in human factors studies. Integrated the two tools into VSCS and re-designed them based on the new requirements.
5) Developed the chat panel for the VSCS and re-designed it based on the new requirements.
6) Developed a simple joystick reading program. This program can read the joystick inputs, axes, rotations, and button clicks. The joystick will be used to control the camera of the UAVs in the VSCS.
7) Created and implemented a message sending test tool for the VSCS. This sends messages to the control station which then sends out the message to the appropriate tool. The tool intercepts the message and performs whatever task is needed.
8) Ported the GITZ (Get in the Zone) algorithm that was written for Linux machines to be able to run on Windows. GITZ is the main focus of the Vigilant Spirit-1 study. GITZ was designed to be easier for UAV operators in maintaining situational awareness while switching between multiple UAVs.
9) Developed a Synthetic Vision overlay tool to be implemented in the VSCS video tool. This tool will draw synthetic objects such as wire boxes or flags to mark points of interest in the video display.
10) Completed the messaging code and fully tested all the message formats.
11) Checked every tool and made sure that all the data that should be recorded during a trial will be recorded.
12) Continued working on the GITZ transition code which still needs to be integrated with VSCS and made sure that it works correctly with the new MetaVR database.
13) Tested the joystick an sensor model in the new database.
14) Checked that the script processor is able to dynamically add vehicle models to the database.
15) Determine how to implement event based actions in the script processor. This allows specific events to be triggered after certain conditions are met instead of being just time based.
16) Re-designed the Synthetic Vision overlay Plug-in tool for the VSCS and verify that objects are being drawn correctly at the right locations in the new database.
17) Finished the GITZ transition code and made sure it is fully integrated with VSCS and works correctly with the new database.
18) Develop additional tools for the Vigilant Spirit testbed as needed.
19) Joystick tested on the sensor model in the new database. Human Factors came up with some improvements in joystick slewing that needs to be implemented.
20) Tested the script processor and verified that it can dynamically add vehicle models to the database.
21) Added a black and white OpenGL shader to the HUD. This changes the video feed to black and white during a GITZ transition.
22) Completed development of min-GITZ Study #1, in which 4 different camera “fly-in” concepts and 3 “fly-in” timings are compared.
23) Updated mini-GITZ-2 proposal and helped with development of transitions and scenarios.
24) Laser designation was added to the Vigilant Spirit Simulation along with event based scripting in the script processor. The script can tell the simulation to wait for a laser designation event to occur before proceeding.
25) Camera elevation angle problems in the vehicle aero model were fixed. Numerous other bugs were fixed in the Vigilant Spirit testbed software.

26) Participated as a subject in the Mini-GITZ 2 study. The results of this study will be used to improve how the GITZ algorithm will function in the Vigilant Spirit-1 GITZ study by selecting the best transition.
27) Updated some of the SIRUS computers so that they can be connected to the Scientific Network.
28) Developed design, implemented scenarios and paperwork, and collected data on 6 subjects for mini-GITZ-2 study.
29) Finished paper on the GITZ development observations from inception to mini-GITZ-1 to mini-GITZ-2 and beyond.
30) Supported GITZ-1 study including script generation, documentation of GITZ development, work with outside resources on new GITZ ideas, work-out "TBDs" on design document, track and report development progress, run subjects.
31) Completed the GITZ-1 testbed software check with Human Factors.
32) Developed and tested training and preliminary script files for experimental trials.
33) Finished writing scripts and configuration files for the GITZ-1 study.
34) Completed data collection.
35) Analysis of data and reported on-going results.
36) Completed data analysis, presented results, wrote paper.
37) Re-designed the chat server and minor VSCS tools.

Predator Mapping Display:

1) Collected and summarized TSD information from pilot and sensor operators from past interviews
2) Met with Predator SPO and General Atomics representatives on Predator TSD design
3) Develop a deliverable SME TSD comments table
4) ACO Tool GUI changes; linked chat to the ACO for requesting airspace.
5) Modified Vigilant Spirit code to support a Glyph study. Wrote software, scripts, and data collection code for the study.
6) Implemented OpenGL Tessallation to help draw concave polygons on the TSD correctly.
7) Started implementing Terrain Avoidance Shading for the TSD.
8) Wrote questions on Airspace, terrain shading, ACO, and killboxes for Fargo Predator operators.

AvantGuard:

1) Met with GamesThat Work learning the software and discussing ideas
2) Worked with new releases of software and reported bugs and issues
3) Reviewed scenario development walk-through document.
4) Reviewed user guides.

AFRL/RHCP

See Addendum 1 and 2

AFRL/RHCV

Informational Display Optimization Laboratory: Human factors research looking at information display optimization. Multiple tasks are included in this project. They include display and night vision device evaluation, e-chart (map) evaluation, the role of bandwidth and nose on

display quality, and how to display information in a manner that improves the user's comprehension and ability to perform a required task.

<u>Task 1: Visualization-Commanders Predictive Environment:</u>

1) Support the customers in day to day discussions.
2) Literature searches are on-going in the areas of uncertainty portrayal, the effectiveness of glyphs and animation in portraying information to users and information theory.
3) Software is being developed to perform experiments designed to evaluate the effectiveness of various information portrayal techniques.
4) Software development and software testing was completed for evaluating the efficiency of information transfer for multi-dimensional Mil TD 2525B symbology. The software was evaluated for the temporal display characteristics to assure that stimuli were displayed for the proper amount of time. Mil STD 2525B symbology sets to be used in experiments were constructed. Experiments are being run currently to evaluate information transfer characteristics of Mil STD 2525B symbology.

<u>Task 2: Evaluation of Short Wave Infrared Sensor (SWIR):</u>

1) Evaluation of SWIR sensor.

General Laboratory Support: Two tasks are being conducted: 1) Digitally Enhanced Video Devices, and 2) Spectral Photometric, Optical and Acuity Evaluations of Electro-Optical Devices.

<u>Task 1: Digitally Enhanced Video Devices:</u>

1) A terrain board was positioned for digital photographs of stimuli to be used in an experimental set-up. Assisted in the set-up , placement of stimuli, and photographs taken. Photographs taken were forwarded to the customer.
2) Photographed Scud Launcher Model in TIFF mode at various angles.
3) A small IR Terrain Board was incorporated into the lab. It was indicated in 5 degree steps and a rotary table fixed to the bottom for ease of rotation about the center axis.
4) Preliminary photos were taken at various angles for analysis and experimental planning.
5) Four Clear P-55 Type visors were prepared and delivered for shipment as part of on-going coating evaluations.
6) The IR terrain board was mounted to an extruded aluminum frame for ease of rotation and a fixed 45 degree orientation elevation. The base was marked in 5 degree increments and a pointer installed for accuracy of indication.
7) Photos were taken of the terrain board for initial evaluation of stimulus placement for experiments.
8) The fused video set-up was moved and oriented to take initial SWIR, FLIR, and NVG images for a baseline.
9) The ITT Intensified Camera was hooked to a Scope to determine if there was an output signal. The manufacturer was contacted about an evaluation of the camera and the process of returning for evaluation was started. It was shipped back to the manufacturer for evaluation and repair estimation.
10) Spectral reflectance scans of various landscape items were downloaded and forwarded for evaluation of NVG and FLIR compatibility to assist in the determination of terrain board landscape.
11) The terrain board on the movement table was switched from the Fall Military Base scene to the Desert Terrain scene.

12) Accomplished spectral scans of scenic materials for evaluation of reflectance compatibility between Visible, FLIR, SWIR, and Night Vision Cameras.
13) Assisted in the gathering of images for DEVD off the terrain board for use in study.
14) Photo documentation of QED Target System.
15)

<u>Task 2: Spectral Photometric, Optical & Acuity Performance of Electro-Optical Devices:</u>

1) Assistance was given in assembly of landing light set up for demonstration to visiting AF personnel.
2) Revised experimental set up to evaluate the optical eye box of PNVG eyepieces. The set-up was use to map the eye box of each eyepiece by translation of a diopter scope along the horizontal and vertical axis until an acuity target lost sharpness. The mounting system for the Diopter Scope was revised to stabilize it for the revised method of measuring the eye box utilizing a 5mm aperture placed over the objective of the diopter scope.
3) Assistance was given in modification of the lighting set-up to be utilized in the evaluation of the "Day-Vision" PNVG. The set-up will be used to evaluate color vision accuracy and visual acuity through the optical system.
4) Developed an experimental set-up to evaluate a LASER detection device using in-house optical hardware. The set-up allowed for a (360) degree azimuth rotation and up to a (30) degree positive elevation adjustment. Data points were taken in (10) degree increments until a non-sign condition was located, then, data points were taken in (1) degree increments.
5) Spectral scans were taken of various materials in support of the SWIRR Camera effort. The scans were used to determine the spectral reflectance of the components photographed in the video sequences.
6) Evaluated Eyepiece diopter settings and measured Eye Box size for PNVG Devices. Performed general Quality Assurance evaluation and documented occlusions according to specifications.
7) Photographed Day-Lite goggles for customers' use in TR. Photographed proto-type filter system for Night Vision Systems. The photographs will be used for SPIE and TR papers.
8) The IR radar detector evaluation was revised and several readings were taken for angles previously omitted, ND filters used and LASER GUN power evaluated using the IL-1700 to determine the power projected on the detector at the test distance. Photos were taken for documentation of the experimental set-up and use in a Technical Report.
9) Evaluated resolution problem on several pair of devices by configuring the lab for 6 different illumination levels and 4 variations of target stimuli.
10) The IR radar detector evaluation was revised and several photos were taken of the internal workings of the commercial detector.
11) A total of ten intensifier tubes were evaluated for gain utilizing the ANV-120.
12) Off-axis measurements of luminance levels and failure to verify the background (Target area) was set for (White) explained the order of magnitude change of test luminance levels.
13) The Pritchard 1980-A was submitted to PMEL for Calibration/Repair. The EG&G Lamp Source was also submitted for Calibration/Certification.
14) The Hoffman Engineering ANV-120 ANVIS Goggle Gain Test System was prepared and shipped to PMEL for Calibration/Certification. It was returned as a "USER CAL" Device. It was shipped to the OEM for lamp replacement, calibration and certification to a NIST traceable standard.

15) The Hoffman LS-65B returned from PMEL was repaired and checked for operation. The read-out was determined to be 4 percent low checking against a Minolta Spotmeter. LS-65B read-out read the same as the Minolta Spotmeter Readout.
16) Photographed proto-type variable transmissive visors and displays.

ROADWIT RHCV: Several projects fall under ROADWIT RHCV: General Labor, Helmet Tracker Requirements Determination, Digitally Enhanced Vision, and Complex Information Display Optimization. The General Labor charges relate to General Support and Night Vision studies being conducted. The Digitally Enhanced Vision is investigating the use of a multi-spectral camera system to enhance target detection.

Digitally Enhanced Vision:

1) General Support
   a) Task 1: Terrain board Target Photos
      i) The photographic requirements were decided on after viewing photos using manual settings with different lighting conditions.
      ii) A cloud shadow shape was produced and several photos were taken using the shadow and three camera orientations. The customer approved this method for shooting a larger series of photos using two shutter speeds for each target quadrant and three camera orientations for a total of 480 photos. The large series of photos was finished. The photos were documented in several Excel spreadsheets and the full series of good photos were placed in a single folder.
      iii) The photographic requirements were specified by another customer. Six photos were taken using manual settings. A cloud shadow was placed in the frame along with two targets. The series of photos included one location of the 16 pie sectors (315 degrees left tilt) used earlier and several new azimuth locations.
      iv) The terrain board computer was tested to see if the updated patches caused operating problems with the computer.
      v) Six wargaming models were assembled.
      vi) Several photos were taken of the FR8 model (Renault R39 Cavalry (light) Tank w/37mm SA 38 Gun). Additional FR8 target test photos were taken utilizing a cloud shadow and manual camera settings.
      vii) Photos were taken of a Scud and an F-15 on the Summer Air Base Terrain board.
      viii) The old terrain board files were reviewed. Manufacturers of the proof-board were located. Samples of the polyurethane board were requested from Goldenwest Manufacturing located in California. From the samples received, the customer chose 10D density proof board for the terrain material.
      ix) Design of the 5 foot terrain board mounting base was started. The boards framework will be designed utilizing 8020 1010 and 1020 materials, The customer required quantity 2, 5 foot terrain boards, one at 700 scale and the other at 285 scale.
      x) The 5 foot Terrain Board Rotary Movement System design was finished
      xi) Several lay-outs utilizing AutoSketch were created to get some idea of how to arrange the landscaping for the forest scene.
      xii) After comparing the scanning data on four different model tree configurations, the yellow and red fall clump foliage tested better from the 700 to 1100 nm spectrum than the medium green clump foliage with T43 yellow grass sprinkles as the tree foliage. Decision to go with the Woodland Scenic FC186 Red Fall Foliage material for the

Deciduous Trees was made. The red foliage was easier to match with the monochrome visible camera.

xiii) Ground foliage in the form of small rocks was added to the terrain board. Small pins were attached to the rocks so they can be placed anywhere on the board.

xiv) The Variable Transmissive Test Cell (VTTC) Visor Systems were photo documented and shipped back to the manufacturer at the request of the customer.

xv) Photo documentation of the QED Target System.

xvi) Small buildings, terrain and landscaping materials have been ordered for the new desert terrain board.

xvii) Preliminary lighting/experimental setup was configured for blooming evaluation of night vision devices.

xviii) After initial testing of blooming/halo lighting the configuration was changed to utilized over-head lighting with variable control to provide the overcast starlight condition at the target and the full length of the test lane.

xix) A platform was built, pre-determined angles marked on the platform, and indicator adapted to subject chair to reference the angles at the testing distances.

xx) The 810nm LEDs arrived and were adapted to the target board. The several power levels were evaluated. The experimental level will be picked from the levels evaluated.

b) Task 2: Runway Lighting (EALS)

i) New 8” x 2” wide caster wheels were specified and ordered for the 4KW CCR cart. The four new 8” pneumatic tires were received.

ii) An 8” tire mounting plate design was completed.

iii) A new PRAMAC power generator was received and the wheel kit was installed.

iv) The packing was completed for Team Patriot.

v) The EALS truck returned from Team Patriot and the lighting system was demonstrated at the HE Open House.

c) Task 3: Multi-Spectral Camera

i) An old synthetic night vision color system platform was located and the night vision cameras were removed from the hand held mount, so the SWIR camera could be attached to the hand held mount. A video cable for the WIR camera and the old system’s LCD monitor was fabricated.

ii) A camcorder with external video recording was recommended to the customer as a recording and display device. Permission was granted to use the Sony Handycam as a recording and display device for the SWIR camera system. Mounting of the Sony Handycam and the SWIR camera on the old synthetic night vision color system’s platform was completed.

iii) A design concept was completed for the new terrain board manual movement system. A requisition and cost estimate was generated for the movement system.

iv) Started setting up the new ITT NVG camera for testing. The auto-iris lens was wired. Testing of the new ITT NVG camera with an auto-iris lens was completed. The camera was demonstrated to the customer.

v) A B/W video camera was located for testing the Nomad HMD system.

vi) Three B/W cameras were located to simulate the video of the multi-spectral cameras. The actual cameras are difficult to use within an artificially lighted room, when trying to troubleshoot the computer system.

vii) The SWIR camera was set-up to test the frame grabber installation.

viii) Briefed on how to operate the camera system software by the software programmer for the multi-spectral camera system.

ix) A demonstration utilizing a resistor as a thermal target source for the LWIR camera was demonstrated to the customer. The customer prefers oven heated to resistive heated model targets.

x) A basic sketch was created of a thermal tri-bar resolution chart for a LWIR camera.

xi) A 2 foot x 2 foot resolution chart was created to fuse the camera images together (3/8 dots 1.25" separation). The camera software required a large target and area, so a 4 foot x 2 ½ foot resolution charge was created to fuse the camera images together with (12.5" Dots 7" separation).

xii) Thirty five thermal points (resistors) were added to the resolution chart so all four cameras can be fused together (Visible, Night Vision, Sort Wave IR, and Long Wave IR).

xiii) Customer requested a multi-camera mounting system be designed. The system will hold 4 cameras. The 4 video cameras are a long wave IR camera, a short wave IR camera, a night vision camera, and a visible camera.

xiv) Widgets completed the fabrication of the multi-camera mount system. Four cameras were removed from the optical bench set-up and placed into the multi-camera mount system. The portable system was mounted to the optical bench for testing.

xv) Customer requested information on small gyro-stabilizers for the portable multi-camera system. A small gyro-stabilizer system was located that would work with the new system. A large variety of camera shoulder stabilizer assemblies were located. The new multi-camera mount system would attach to the shoulder assemblies for portable recording.

xvi) A 12VDC to 120VAC power inverter system was drawn. The design included all the current equipment used for the semi-portable MS camera system. The power consumption was approximately 350 watts.

xvii) Two air deflectors were designed and fabricated for the MS camera computer expansion chassis. The Matrox video cards were reaching their temperature limits. The deflectors lowered the operating temperature of the video cards. Any further reductions need to be done with the increase in air flow.

xviii) Removed the MS camera system from the optical bench and placed it on a portable cart. Three different scenes were taken with all four cameras.

xix) Customer requested a project to develop a combined thermal and visible resolution target for the Multi-Spectral Camera System. The properties of an aluminum target were tested. The aluminum target needs to have a radiator surface either of anodizing or flat black paint.

xx) Assembly and fabrication of the Landolt C Target controller completed. Design and fabrication of the target 8020 frame and stand completed. The controller program has been written and the de-bugging started.

xxi) The controller program has been written and the de-bugging has been completed. The Peltier Heat Pump didn't perform very well. The heat pump was replaced with 5 power resistor. A test thermal plate has been fabricated and a second aluminum plate was added with one side painted flat black to test the new controller and the thermal radiance of the plates. The program was modified to incorporate the new heat source. The controller and its program have been completed.

xxii) A manufacturing problem developed with the QED thermal panels. The panes were re-made. The QED system was tested and a shadow was cast by side lighting because of the target panels recess. The target surface plates need to be modified by increasing the depth of the recess pocket.

xxiii) The QED target system electronic test has been completed and data collection started.

xxiv) A user manual for the QED system has been completed.

xxv) A monitor mount utilizing 8020 material was fabricated for an LCD monitor.

xxvi) A motorized QED Target Stand design was started. Remote controlled motorized lateral adjustment will be incorporated into the design. The vertical adjustment will be manual, but it can be motorized in the future.

xxvii) The QED target system continued being functionally tested.

xxviii) A design for the QED target stand remote control motorized horizontal and vertical axis has been completed.

xxix) The Multi-Spectrum Camera System was utilized to collect video images for Dr. Reppenger and WSU's Dr. Pink.

xxx) A video problem was detected in the SWIR camera. The camera was shipped back to the manufacturer. Sensors Unlimited found a problem with the power supply. They replaced and are shipping back.

xxxi) The customer wanted to prototype a new device. The device would combine two video images. The device consists of a beamsplitter and a mirror with appropriate wavelength properties. The thermal camera would be in the position that gets the image as it bounces off of both the beamsplitter and the mirror. The visible camera, or the NVG wavelength camera, or the SWIR would the "other" camera mounted to receive the straight through light of the beamsplitter. The "other" camera need to be mounted so that it can be slightly adjusted in Az, EL, and Roll to align the two images.

xxxii) A prototype breadboard was assembled from in-house materials on a portable optical table so the breadboard can be taken outside for image collection.

d) Task 4: General Support for the Digitally Enhanced Vision Lab, the Night Vision Operations Lab as well as the Windscreen Lab

i) The eye-box on six pair of PNVGs was measured and evaluated using the set up that involves translating a diopter scope along the horizontal and vertical axis until the target is just out of focus. The eye-box on a single F4949 tube was also measured and evaluated.

ii) Visual acuity data was collected. Acuity measurements were made using F4949 NVGs, both with and without the filters as well as at two light levels.

iii) Assistance provided with measuring and evaluating the eye-box on several pair of PNVGs. This measurement technique involves using the set-up that involves translating a diopter scope along the horizontal and vertical axis until the target is just out of focus.

iv) Attended a meeting to discuss visual anomalies that were detected in several pairs of PNVGs during the eye-box evaluation.

v) The small COTS IR terrain board was mounted to allow for a 360 degree rotation capability for digital image stimulus generation. Images will be generated using four different types of cameras (IR, SWIR, Visible and Thermal) and will be used in psychophysical evaluations of different image enhancing algorithms.

vi) Editing was completed for two papers/briefings to be presented at the ROT Human Factors & Medicine Panel, HFM-141 symposium on "Human Factors and Medical Aspects of Day/Night All Weather Operations: Current Issues and Future Challenges" [18]in Heraklion, Greece. The second paper will be presented at SPIE in Orlando, FL.

[18] RTO Human Factors and Medicine Panel, HFM-141 Symposium "Human Factors and Medical Aspects of Day/Night All Weather Operations: Current Issues and Future Challenges"

e) Task 5: DVED Lab
   i) Sequence files were generated to be used in a small magnitude estimation pilot study investigating non-degraded images as function of contrast reduction. Data sheets were completed. Data was collected from four subjects. Each subject gave a subjective numerical rating when comparing a degraded image to a standard (non-degraded) image. The task was repeated four times.
   ii) Met with the customer and the statistician to discuss the experimental design for several studies investigating the performance of multiple algorithms. A draft protocol was written in preparation for the objective assessment of images for DVED.
   iii) Assistance was provided in editing and formatting a paper entitled “A Unified Taxonomical Approach to the Laboratory Assessment of Visionic Devices”.

2) Helmet Tracking Requirement:
   a) General Support
      i) The design and fabrication were completed for mounting an optical lens holder between a subject and a pair of NVGs for Col. Baldwin.
   b) NVG Adjustment Methods, Eyepiece Focus Settings, and Vision Study: Experiments will be performed which will investigate vision as a function of eyepiece focus settings to include visual acuity, accommodation, clarity and comfort. Eyepiece focus settings as a function of adjustment method, to include monocular vs. binocular techniques, use of the ANV-20/20 vs. distant targets and use of lens bars vs. snap-on lenses will be investigated.
      i) Data collected to investigate two methods of determining the required snap on lenses for the user to wear on a pair of PNVGs. The current method of snap on lens selection that is outlined in the PNVG T.O. was compared to another method where the subject used a lens bar that contained several lenses in .5 diopter increments to determine lens selection.
      ii) Data collected investigating two focusing techniques: focusing F4949 NVGs using the Hoffman 20/20 and focusing using a point source of light which is simulating a star in the night sky.
      iii) Meeting held to discuss the experimental design of the next phase of the NVG Adjustment Methods, Eyepiece Focus Settings, and Vision Study. This phase will involve measuring the subject’s visual acuity with different trial lenses placed in front of the eye, simulating different NVG eyepiece focus settings. Data sheets were prepared.
      iv) Data collected from one subject investigating two focusing techniques: focusing F4949 NVGs using the Hoffman 20/20 and focusing using a point source of light which is simulating a star in the night sky.
      v) Visual acuity data was collected on thirteen pairs of NVGs using the ANV 126 test kit as well as the Hoffman 20/20. Visual acuity was recorded on all four channels of the PNVGs.
   c) Adjustable Brightness Control (ABC) NVG Study: The Adjustable Brightness Control Night Vision Goggles enable the pilot to increase the NVG output luminance to potentially enhance visual acuity, while viewing through goggles. As pilots become adapted to these brighter than usual luminance levels, their ability to read their cockpit instruments while looking under the NVGs may be degraded. In addition, the NVG output luminance levels may affect the time required for the pilot to re-adapt to their cockpit lighting environment, and to allow him to identify and discern necessary information on approach while viewing outside the cockpit window without goggles. The Performance Assessment

of the Adjustable Brightness Control Night Vision Goggles Study is investigating visual acuity data at 3 different NVG output luminance levels as well as the time required for the human visual system to recover after adapting to the NVG output luminance.

i) A small study to verify dark adaptation levels using 2 subjects was completed.

ii) Both the positive and negative contrast of the target displayed on the Micron computer was measured, with and without NVGs.

iii) Attended meeting with the statistician to discuss the preliminary statistical analysis of the data collected from the ABC Study.

3) Digitally Enhanced Vision: The Digital Visual Enhancement Device is being designed to replace the current NVG image intensifier tube with a solid-state digital device that will contain built-in computational capabilities that will allow agile, real-time image enhancement. A series of vision enhancement algorithm studies are being conducted to assess the quality of image enhancement algorithms by comparing target detection with and without the image processing. Support is also provided to the FAA, ASTM, and the RTO working groups under this work unit.

a) All IRB training modules completed. The certificate of completion was sent to the IRB coordinator.

b) Public Affairs clearance dates were researched and recorded to allow for publications to be included in the RHC web-site.

c) Assistance was provided to the customer in preparing a briefing to the American Society for Testing Materials conference.

d) The protocol for The Effects of Image-enhancing Algorithms on Visual Performance study was received with editorial comments from the IRB and the Base legal department. The necessary changes were completed and the protocol was submitted for final IRB approval.

e) Telecon with UDRI and RHCV personnel regarding RHCV's support in evaluating prismatic deviation, optical/refractive power and distortion of visors from the Joint Helmet Mounted Cueing system office.

f) Meeting held with the Principal Investigator to review the experimental design of The Effects of Image-enhancing Algorithms on Visual Performance study.

g) The Effects of Image-enhancing Algorithms on Visual Performance study was completed. The data was forwarded to the customer as well as the statistician for further analysis.

i) NVG Adjustment Methods, Eyepiece Focus Settings, and Vision Study: Experiments will be performed which will investigate vision as a function of eyepiece focus settings to include visual acuity, accommodation, clarity and comfort. Eyepiece focus settings as a function of adjustment method, to include monocular vs. binocular techniques, use of the ANV-20/20 vs. distant targets and use of lens bars vs. snap-on lenses will be investigated.

(1) Data collected to investigate two methods of determining the required snap on lenses for the user to wear on a pair of PNVGs. The current method of snap on lens selection that is outlined in the PNVG T.O. was compared to another method where the subject used a lens bar that contained several lenses in .5 diopter increments to determine lens selection.

(2) Data collected investigating two focusing techniques: focusing F4949 NVGs using the Hoffman 20/20 and focusing using a point source of light which is simulating a star in the night sky.

(3) Meeting held to discuss the experimental design of the next phase of the NVG Adjustment Methods, Eyepiece Focus Settings, and Vision Study. This phase will involve measuring the subject's visual acuity with different trial lenses placed

in front of the eye, simulating different NVG eyepiece focus settings. Data sheets were prepared.

(4) Phase 3 of the NVG Focus study began. Visual acuity data was collected. This phase involves measuring the subject's visual acuity using a pair of F4949 NVGs set to 0.0 diopters. Ophthalmic trial lenses are placed between the subject's eye and the eyepiece of the goggles. Visual acuity data was collected both monocularly and binocularly with eight different trial lens conditions.

(5) Phase 3 of the NVG Focus was completed. Visual acuity data was collected from a total of eight subjects.

h) Transmissivity, haze, multiple imaging, and internal reflection data were measured of a variable transmittance visor test cell.

i) Photo documentation of Variable Transmissive Visor. Transmission and haze data was collected on the same system. The RaDOMA Spectradiometer, Gardner Hazemeter, and Minolta Spotmeter were utilized for data collection.

j) Paper, entitled "Quad-emissive Display for Multi-Spectral Sensor Analyses",[19] was edited and formatted for submission to Fusion 2008 conference.

k) Paper, entitled "Dynamic Stimulus Enhancement with Gabor-based Filtered Images", was edited and formatted for submission to SPIE.

4) **CNRC (Canadian National Research Center) Helmet Mounted Photometer and NVG Ambient Illumination Tester**

a) The CNRC is interested in using the Helmet Mounted Photometer (HMP) and NVG Ambient Illumination Tester (AIT) to take Day and Night ambient light level readings from a helicopter platform.

b) Re-programming of the Helmet Mounted Photometer (HMP) software was started. The HMP is being programmed to output Real-Time luminance values from the HMP photometer head and NVG Ambient Illumination Tester (AIT) voltage output to a computer.

c) The field of view changes when the gain of the TSL230 Light to frequency IC changes. A user option was added to the program to fix the gain to 10.

d) User manual for the CNRC HMP-AIT combine system is finished and submitted for editing.

e) Arrangements are being made to ship the Head Mounted Photometer and the NVG Ambient Illumination Tester to the CNRC.

f) CNRC sent AFRL/RHCV an Illuminator device to be calibrated.

g) The illumination device output was prepared for shipment after completing the testing and documentation of the device.

5) **JIEDDO Support:** The specific goal of the research is to quantitatively measure various aspects of vision function and compare them to the speed and accuracy of target detection. A field study as well as an in-house study will b e conducted to determine if there is a correlation between visual function ability and the speed and accuracy of IED target detection.

a) Meeting held to discuss the design of the in-house study as well s areas of responsibility and data to be collected for the field study.

b) Current information/briefings regarding IEDs were reviewed.

[19] Fusion 2008 Conference "Quad-emissive Display for Multi-Spectral Sensor Analysis"

c) A draft protocol was prepared for the in-house study to be conducted and submitted to the customer for review.
d) Visual function tests/metrics were reviewed. These tests include: 1) visual acuity; 2) contract sensitivity; 3) color vision; 4) depth perception; and 5) refractive error.
e) Color vision data using the FM-100 Hue test collected at Ft. Campbell was reduced and analyzed using only two of the four boxes of the set.
f) Attended an organization meeting at the U.S. Army Research Institute for the Behavioral and Social Sciences Simulator Systems Research Unit in preparation of field data collection.
g) The data sheets for the field data collection was modified to reflect the changes to the vision metrics that will be evaluated.
h) Preparations for the data collection in Twentynine Palms, CA continued. Equipment, data sheets, informed consent documents and supplies were assembled and prepared. Preparations for the data collection at Ft. Dix, NJ, continued.
i) Vision data collected on approximately 85 subjects at Ft. Dix Army Post. Data entered and preliminary statistical analyses performed.
j) Vision data collected at Ft. Sill on 40 subjects entered and preliminary statistical analyses performed.
k) Vision data collect at Ft. Leonard Wood entered and preliminary statistical analyses performed.

6) **Strike Info Displays:** Research visual displays for Command & Control applications
  a) Graphical Primitive Development for Volumetric Display
    i) Developing graphical primitives (shapes, colors, styles, etc.) for viewing on the Perspecta 3D volumetric display
    ii) Finished data analysis.
    iii) Final report completion, summarized results of extensive data collection, literature review, etc., regarding graphical primitives
    iv) George Reis orally presented final report at SPIE D&S 2008
    v) Developed specific requirements of virtual reality technologies for military applications.
    vi) Developed a visual search experiment for investigating the amount of stereoscopic disparity that is necessary to induce the “pop-out” effect.
    vii) Developed a set of design guidelines for cyberspace development that will enable effective navigation; comparative animal navigations methods will be discussed.
    viii) Developing specific requirements of virtual reality technologies for military applications.
  b) 3D Display Metric Development:
    i) Developing useful measurements and specifications (metrics) for comparing and contrasting a variety of 3D display technologies.
    ii) Final report completion, summarized literature review, thoughts on the review, and some spectra-radiometric data collected concerning objective measurements of 3D displays
    iii) Dr. Paul Havig orally presented final report at SPIE D&S 2008
  c) Tangible User Interface Evaluations:
    i) Developing and evaluating tangible user interfaces for interacting with three-dimensional data sets
    ii) Final report completed, summarized extensive literature review and subjective evaluations of in-house tangible user interface technologies.

d) Network Visualizations – creation and evaluation
   i) Developing and evaluating network visualizations for cyberspace situation awareness

7) **Desert Terrain Board:**
   a) Small buildings, terrain and landscaping materials have been prepared for mounting to the terrain board by gluing pins to them.
   b) An external parallel ZIP drive as installed on the Radoma spectrometer's original computer. This allows programs and collected data to be transferred to other computers, which in turn makes the Radoma a viable system. Spectral scans were taken of Quikrete medium sand for the terrain board. The Quickrete medium sand was glued to a sample board using Woodland Scenic Landscape Cement. The Quikrete sand will work for the desert terrain board if a good method for application can be determined.
   c) Several paint samples were scanned.
   d) All the small buildings have been prepared for mounting to the terrain board. Testers F414302 paint matches the technical report dry sand measurements within 10 percent from 400nm through 1350nm.
   e) Calculations completed on the amount of paint required to cover the terrain board.
   f)
8) **General Support:**
   a) Several meetings attended to discuss the test and evaluation of the SU640 SWIR camera. The spectral response measurement procedure was documented and submitted to the POC of the enhanced SU640 SWIR cameras. The SU640 SWIR camera spectral response and bad pixels were tested. Additional testing on SU640 SWIR cameras will be tested after their flight test.
   b) A spectral response measurement was completed on the enhanced SU640 SWIR camera.
9) **Quad Sensor Array (QSA) and Quad Emissive Display (QED) Testing**
   a) Testing continues on the QED and QSA devices to prove their capabilities. Two tests were conducted on each of the four cameras to determine the effects of the frame grabber capture on each of the cameras' video signal. A multitude of images were captured at a 20 meter distance. A total of 224 images were captured for a third test. The images were used to test an automatic Landolt C orientation determination software program developed in-house.
   b) A new design was finished for a portable Quad Emissive Display (PQED). The unit would be about 5.75" square by .25 to 3.5 inches and the mount design would attach to a tripod. The unit's vertical target surface would rotate freely or lock at every 90 degrees. Testing continues on the QED and QSA devices to prove their capabilities. A total of 524 images were captured for an additional test. The images were used to test an automatic Landolt C orientation detection software program developed in-house.
   c) During testing of the QED and QWSA devices, the LWIR camera images became unreadable because of the sensitivity to where the gap falls with respect to the CCD pixels. The target was positioned both vertical and horizontal to produce the best image before capture.
10) **PCALS –** System shipped
11) **Micro Vision**
   a) The new Micro Vision HMD system was received and tested. The red color of the image was lost within the first ten minutes and all three colors were gone within an hour after testing begun. The system was returned to the manufacturer. The Red Laser fault was caused by a broken connection from the Laser to the Flex cable. T he Red Laser anode

became disconnected. The system safety circuit shut down due to no laser feedback signal. The current ASIC chip set required an impedance matched cable to support the system electronics partitioning. The cable was only used to support this demonstrator while waiting for the development of a new ASIC chip set. Plan is to have the new chip set and eliminating the cable for the DARPA Ultra-Vis system.

b) The Microvision HMD system was returned to the manufacturer. The system was powered up and the display was observed for about 30 minutes. The second time it was powered up and operated for about 45 minutes, the green laser stopped working. The system was powered up the third time and the display head started to make noise, so it was immediately powered down. The manufacturer stated this nose was normal during a calibration procedure and it would take approximately 15 minutes. After the calibration procedure finished, the green laser started functioning again. Photographs of the display were taken of a computer desktop with 6 different background colors to show the readability of the display. Photos were taken with six different desktop colors, at three different brightness levels, and two different lighting conditions. The lights on condition: a white piece of paper was placed in front of the HMD device to capture some of the room brightness. The images captured showed a little more icon blooming, than in the real observation of the display. The icons and the text were still not legible in the HMD device with any background color or brightness. A Nikon CoolPix 8800 digital camera was used to capture the images. The camera was set to manual focus and fixed aperture. The optimum shutter speed was set by the camera, which produced the mild icon blooming effect.

RHCV (6000)

Tracker Development:

1) Performed final install of tracker and skid into aircraft
2) Performed first and second flight of Ascension tracker
3) Trained other participants on the operation of the spectrometer
4) Performed Gimbal repeatability tests using laser
5) Built and tested a timing circuit for the tracker
6) Re-installation of optical tracker skid assembly in aircraft
7) Second test flight in Cleveland successful
8) Continued technical support for the testing on NASA aircraft
9) Updated analysis report for the Gimbal repeatability
10) Developed software for the USB timer circuit
11) Assembled a 19” rack panel for head tracker test equipment
12) Developed software to post analyze spectrometer data collected during flight test
13) Wrote instruction manual for operation of spectrometer and trigger circuit
14) Evaluated methodology used to collect repeatability data for the Gimbal
15) Collected preliminary repeatability data to determine the Gimbal’s contribution to the overall accuracy of the tracker system
16) Modified and tested a timing circuit for the encoder to interface card
17) Fabricated accelerometer cable for HBM system
18) Completed fabrication of USB Timer circuit board
19) Laser tracker screen set-up completed
20) Completed build of several prototype circuits to evaluate the USB timer circuit

General Support:

1) Working issue of GFE equipment used for CATS
2) Installed video cards for use in wireless HVI SBIR
3) X-Ray of HALM missile to determine safety
4) Components for the 3 axis motion controller and table received
5) Component assembly and wiring of the XYZ Motion Controller chassis begun
6) Evaluated the performance of the SWIR camera in a camouflaged environment under various lighting conditions
7) Evaluated IPNVG equipment
8) Trouble-shot LPIDS and suitcase PIDS and made necessary repairs
9) Writing SPIE paper
10) Received, installed and tested second high speed serial card.
11) Completed test of software to support wireless HVI test.
12) Reviewed and discussed interconnectivity requirements for BAO equipment when multiple displays and controllers are utilized.
13) Evaluated SH21 IPNVG goggles and diagnosed problems that need repaired by the manufacturer.
14) Developed and tested a sync interface circuit for use in wireless HVI SBIR
15) Researched installation and removal tools available for the Melles Griot optics table
16) Designed and procured a specialized tool for the optics table
17) Completed 5 composite Tool Kits
18) Completed and tested ICUITI interface cable

Information Visualization: Multi-spectral image enhancement and fusion for man-in-the-loop experimentation/evaluation

1) Developed the foundation of a software tool for image and video capture
2) Implemented Multi-scale retinex enhancement and fusion algorithms, some noise removal algorithms and thermal cuing enhancement algorithms.
3) Implemented multiple enhancement utility algorithms in an open, flexible modular architecture.
4) Integrated camera system with newer mechanical mountings.
5) Continued work on real-time algorithm development for natural scene registration.
6) Measuring performance of auto-registration algorithms using FPGAs, modern CPU (Intel Core 2 Duo) with use of MMX, SSE instructions and using nVidias GPU chipsets (CUDA).
7) Obtained Sarnoff algorithms.

NALEP: A study will be conducted to assess night vision goggle damage versus laser hardening.

1) Data collected.

Abbreviations/Acronyms

AFs	Articulatory Features
ASR	Automatic Speech Recognition
CMLLR	Constrained Maximum Likelihood Linear Regression
CMU	Carnegie Mellon University
CSMAPLR	Connstrained Structural Maximum-A-Posteriori Linear Regression
DLIFLC	Defense Language Institute Foreign Language Center
DLL	Dynamic-Link Library
GLOSS	Global Language Online Support System
GMMs	Gaussian Mixture Models
GUIs	Graphical User Interfaces
HMM	Hidden Markov Model
HSMMs	Hidden Semi-Markov Models
HTK	HMM Hidden Toolkit
HTS	HMM Speech Synthesis Toolkit
ILR	Interagency Language Roundtable
IPA	International Phonetic Alphabet
KLT	Karhunen-Loeve Transformation
LASER	Language And Speeech Exploitation Resources
LMs	Language Models
MAP	Media Analysis Plug-ins
MDL	Minimum Description Length
MFCCs	Mel-Frquency Cepstral Coeefficients
MLPs	Multi-Layer Perceptrons
MSD	Multi-Space probability Distribution
MT	Machine Translation
PER	Phoneme Error Rate
PLP	Perceptual Linear Prediction
ROADWIT	Research Operations for Advance Warfighter Interface Technologies
ROVER	Recognizer Output voting Error Reduction
SAT	Speaker Adaptive Training
SCREAM	Speech and Communication Research, Engineering, Analysis, and Modeling
SDK	Software Develoment Kit
SI	Speaker Independent
SPTK	Signal Processing Toolkit
TCP/IP	Transmission Control Protocol/Internet Protocol
TDT4	Topic Detection and Tracking 4
TRANSTAC	Translation System for Tactacle Use
VTLN	Vocal Tract Length Normalization
WERs	Word Error Rates
WSJ0	Wall Street Journal
WSJ1	Wall Street Journal